

School of Economics and Finance  
Queen Mary University of London

# Essays on Human Capital in Brazil

Martin Foureaux Koppensteiner

A thesis submitted in partial fulfilment of the requirements of the Degree of Doctor of  
Philosophy in Economics.

October 2013

### Statement of conjoint work

The third chapter draws on joint work with Marco Manacorda. I was responsible for the data analysis and the writing was shared equally.

To my parents, Traudl and Harald Koppensteiner.

## **Abstract**

This thesis consists of three independent chapters on human capital in Brazil.

Chapter one examines the effect of the introduction of automatic grade promotion on student performance in 1,993 public primary schools in the Brazilian state of Minas Gerais. A difference-in-differences approach that exploits variation over time in the adoption of the policy allows the identification of the treatment effect of automatic promotion. I find a negative and significant effect of about 6% of a standard deviation. I provide evidence from quantile regression DiD estimates consistent with an interpretation of the findings as disincentive effect on student effort associated with the introduction of automatic grade promotion; additional evidence on student and teacher behaviour supports this interpretation.

Chapter two provides a novel way of identifying peer group effects in Brazil. Students in Brazil are typically assigned to classes based on the age ranking in their cohort. I exploit this rule to estimate the effects on maths achievement of being in class with older peers for students in fifth grade using a regression discontinuity design. I provide evidence that heterogeneity in age (and in other characteristics) is an important factor for student performance. Information on teaching practices and student behaviour sheds light on how class heterogeneity may harm learning.

Chapter three uses microdata from Brazilian vital statistics natality and mortality data between 2000 and 2010 to estimate the impact of in-utero exposure to local violence - measured by homicide rates - on birth outcomes. Focusing on small communities, for which it is more credible that local homicide rates reflect actual exposure to violence, the analysis shows that exposure to violence during pregnancy leads to deterioration in birth outcomes: one extra homicide during the first trimester of pregnancy increases the probability of low birthweight by around 6 percent.

## Acknowledgements

I am deeply indebted to my supervisor Marco Manacorda. I benefited greatly from his invaluable advice and support, his constructive criticism and encouragement throughout the PhD. I also would like to thank Francesca Cornaglia and Randi Hjalmarsson for their great support during my PhD.

Each chapter also benefited from comments and feedback received from discussions, and from seminar and conference participants. The first chapter benefited from comments by David Card, Jefferson Mainardes, Imran Rasul and seminar participants at University College London, Queen Mary, the International Conference on the Economics of Education in Zurich, SMYE, and the Congress of the EEA. For the second chapter, I am grateful to Habiba Djebbari, Claudio Ferraz, Giacomo de Giorgi, Lars Lefgren, Barbara Petrangolo, Rodrigo Soares, Petra Todd, Maximo Torero for valuable feedback. I also would like to thank seminar participants at PUC Rio, Queen Mary, Centre for Economic Performance LSE, Alicante, Leicester, ZEW Mannheim, RES Meeting, ESPE, the North American Winter Meeting of the Econometric Society, the EALE/SOLE 3rd International Conference, the 13th IZA Summer School in Labor Economics, the Congress of the EEA and the Conference of the German Economic Association for useful comments. The third chapter benefited greatly from comments by Ana Corbacho, Carlos Scartascini, Rodrigo Soares, and seminar participants at the IDB's first and second seminars "The Costs of Crime and Violence in Latin America and the Caribbean". Financial support from the IDB for the third chapter is very gratefully acknowledged.

I also would like to thank the Secretariat of Education in Minas Gerais, the Brazilian Ministry of Education, the National Institute for Educational Studies and Research (INEP), and the Brazilian Ministry of Health for the data. I am grateful to Juliana Riani, Jorge Rondelli and Dácio Rabello for assistance obtaining the data.

Finally, I would like to thank my wife for her infinite patience and her love and encouragement throughout the PhD.

## Table of contents

<b>Abstract</b>	<b>4</b>
<b>Acknowledgements</b>	<b>5</b>
<b>Table of contents</b>	<b>6</b>
<b>List of tables</b>	<b>9</b>
<b>List of figures</b>	<b>11</b>
<b>Preface</b>	<b>12</b>
<b>1. Automatic grade promotion and student performance</b>	<b>15</b>
1.1 Introduction	15
1.2 The school system in Brazil and Minas Gerais	18
1.3 Data description	19
1.4 The General Education Act of 1996: the case of a quasi-experiment	21
1.4.1 Policy background	21
1.4.2 Assignment to treatment	22
1.5 Empirical strategy	25
1.6 Estimation results	27
1.6.1 Main results	27
1.6.2 Interpretation of the results and the disincentive effect of automatic promotion	29
1.6.2.1 Changes in the student composition	31
1.6.2.2 Introduction of automatic promotion at 2 <sup>nd</sup> grade	35
1.6.2.3 Effect of the policy change on drop-out rates	37
1.6.2.4 Effect of the policy change on school transfer rates	37
1.6.2.5 Systematic test taking behaviour	38
1.6.2.6 Effect of the policy change on class size	38

1.7 Conclusions	39
<b>2. Class Assignment and Peer Group Effects</b>	<b>57</b>
2.1 Introduction	57
2.2 The allocation of students into classes in Minas Gerais	61
2.3 Data and descriptive statistics	62
2.4 Empirical strategy	64
2.5 Empirical results	65
2.6 Tests for non-random sorting	68
2.7 Interpretation of the effects	70
2.7.1 Exogenous peer characteristics and direct peer effects	71
2.7.2 Indirect effects: responses of schools	72
2.7.3 Indirect effects: responses of teachers and students	74
2.7.4 Opening the black-box of the peer-group effect: heterogeneous treatment across schools	77
2.8 Conclusions	80
2.9 Data appendix	93
2.10 Appendix on initial class assignment and class transition	97
2.11 Appendix on selection of schools in the sample	99
<b>3. The Effect of Violence on Birth Outcomes</b>	<b>110</b>
3.1 Introduction	110
3.2 Birth Outcomes and In Utero Experiences: The Effect of Exposure to Violence	113
3.3 Background, Trends, and Data	115
3.3.1 Births and Birth Outcomes	115
3.3.2 Infant Mortality	117

	8
3.3.3 Homicides	117
3.4 Econometric Methodology	120
3.5 Empirical Results	122
3.5.1 Birthweight	122
3.5.2 Additional Outcomes	124
3.5.3 Alternative Definitions of Homicide	125
3.5.4 Heterogeneous Effects by Mother's Education	125
3.5.5 Summary of Findings and Conclusions	126
<b>References</b>	<b>139</b>



## List of tables

1	Test Score Means in Treatment and Control Schools before and after the Adoption in the Treatment Schools	42
2	Main Estimation Results and Sensitivity to Age Controls	42
3	Effect of the Adoption of Automatic Promotion on Student Flows and Class-size	43
4	Estimation Results for Restricted Age Ranges	44
5	Effect of the Adoption of Automatic Promotion on the Socio-economic Composition	45
6	Effect of the Adoption of Automatic Promotion on Participation in PROEB	45
7	Effect of Policy Adoption on Student Net Flow	46
8	Mean Characteristics of Treatment and Control Groups in 2003 and 2006	48
9	Mean Pre-intervention Repetition Rates in 1997	51
10	Sensitivity of Estimates to Individual and Peer Age Controls for Different Age Ranges	53
11	Effect of Policy Adoption on 3rd Grade Repetition Rate for Treatment Schools	54
12	Linear Probability Model of Assignment to Treatment	55
13	Means and Proportions of Student and Teacher Characteristics	83
14	Main Estimation Results	84
15	RD Estimates of Maths Test Scores	85
16	RD Estimates of Predetermined Individual and Family Variables	86
17	Class and Teacher Characteristics	87
18	Response of Teaching Practices to Class Composition	88
19	Heterogeneous Treatment Across Schools	89
20	Means and Proportions School and Headmaster Characteristics	101
22	Socioeconomic Characteristics of Students in Schools in Sample/ not in Sample	102
21	Choice of Class Assignment Rule	103
22	Means of Student Statements on Teaching Practices and Peer Behaviour	106
23	RD Estimates of Predetermined Individual Characteristics of the 2007 Entry Cohort	106
24	Descriptive Statistics	128

25	The Effect of Homicides during Pregnancy on Birthweight by Trimester since Conception – Small Municipalities	130
26	Homicide Rates and Additional Birth Outcomes by Trimester since Conception – Small Municipalities	131
27	The Effect of Homicides during Pregnancy on Birthweight by Trimester since Conception –Alternative Definition of Homicide Rate - Small Municipalities	132
28	Homicide Rates and Additional Birth Outcomes by Trimester since Conception – by Mother’s Education – Small Municipalities	133
29	The Effect of Homicides during Pregnancy on Birthweight by Trimester since Conception – By Municipality Size	134

## List of figures

1	Histogram for the Propensity Score (year 2003)	47
2	Treatment Scheme	47
3	Local Averages and Local Linear Regression of Treatment and Outcome Variable	90
4	Distribution of RD Estimates Across Schools	91
5	Age Distribution in the Cohort	91
6	Age Distribution in Younger Classes	92
7	Age Distribution in Older Classes	92
8	Test for Discontinuity of Individual and Peer Values of Pre-determined Characteristics	107
9	Municipality Characteristics.	135
10	Incidence of Low Birthweight and Homicide Rates across Municipalities	136
11	Effect on Outcomes by Trimester since Conception	137
12	Small Municipality Characteristics	138

## Preface

This thesis contains three independent chapters that are aimed towards contributing to our understanding of human capital accumulation in Brazil. The first two chapters are concerned with Brazilian primary education. The education system in Brazil underwent a major expansion in the last two decades and access to primary and lower secondary education today is almost universal. The rapid expansion of public primary education created challenges for schools and policy makers, including widespread grade retention and large age heterogeneity within student cohorts due to a combination of late enrolment, suspension of studies and grade retention.

Chapter one looks at a policy aimed at reducing grade retention by automatically promoting students from one grade to the other. In the state of Minas Gerais automatic promotion was introduced in the 1,993 public primary schools gradually over time. I employ a difference-in-differences strategy that uses the variation in the timing of the adoption of the policy to estimate the effect of removing the deterrent of grade retention on standardized test scores. I find that the introduction of automatic promotion significantly reduces academic achievement measured by math test scores of 4<sup>th</sup> graders by 6% of a standard deviation. Under plausible assumptions I argue for the interpretation of the results as the disincentive effect of the introduction of automatic promotion. I provide evidence that heterogeneity in age (and in other characteristics) is an important factor for student performance. Information on teaching practices and student behaviour sheds light on how class heterogeneity may harm learning. The chapter fills a gap in the literature by looking at the ex-ante effects of grade retention and helps to explain the persistence of grade repetition in many countries as means of incentivizing students. The findings are also important because they reveal that a large fraction of students is affected by the grade promotion regime, not only students that are actually retained.

The second chapter provides estimates on peer effects using quasi-experimental variation in peer group membership in 5<sup>th</sup> grade classes of primary schools in Brazil. Primary school students are typically allocated to classes based on their relative age within their cohort, to make classes more homogeneous in age. Using the age rank as a continuous assignment variable, this rule creates a discontinuity in the allocation to a class and peer group for students close to the cut-off point. Focusing on schools with two classes per cohort I exploit the discontinuity caused by the assignment mechanism to compare outcomes of students on either side of the cut-off point. I find that marginal students who are assigned to the older classes have maths test scores that are around half of a standard deviation lower than those of students assigned to the younger classes. I provide additional evidence from variation across schools pointing to the importance of the difference in age dispersion between classes for explaining the estimated group effects. Complimentary outcomes from teacher and student questionnaires reveal that the class composition is also associated with behavioural differences of students and teachers in the classroom. The chapter provides a novel way of identifying peer effects using an assignment rule of students into classes that uses relative age of students. The results also contribute to the discussion on streaming of students into classes and schools and to the understanding of the effect of policies that aim at reducing the age variation in a given school cohort.

In the third chapter (joint work with Marco Manacorda) the perspective on human capital changes from primary education to the time in utero. Using a very rich dataset on the universe of births and homicides from vital statistics data over the period 2000-2010, we estimate the effect of in-utero exposure to homicides on a range of birth outcomes in small Brazilian municipalities, for which it is more credible that local homicide rates reflect actual exposure to violence. Identification is based on a difference-in-differences strategy across geographical areas and time. We find a

significant negative effect of exposure to violence during the first trimester on birthweight, which is line with findings on the effect of other stress-related shocks during pregnancy in the literature. We also find significant and large positive effects of homicides on the probability of low birthweight, implying that the effects are particularly pronounced at the bottom tail of the birthweight distribution. The results are largely concentrated among poorly educated mothers. This suggests that violence adds up to the mechanisms that affect the transmission of socioeconomic status between parents and their offspring. We provide evidence that violence in the first trimester of pregnancy affects birth outcomes through reduced gestational length, rather than intrauterine growth retardation.

# 1. Automatic grade promotion and student performance

## 1.1 Introduction

Grade retention, the practice of holding back students in the same grade for an extra year if they fail to achieve promotion requirements – either in the form of a performance measure or in the form of minimum attendance – is used in many developing and in some developed countries. It is particularly widespread and pronounced in African and Latin American countries, where repetition rates are often as high as 30% (UNESCO 2008).<sup>1</sup> Historically grade repetition had a prominent role in Brazil and repetition rates in Brazilian primary schools reached 24% in first grade and 14% in fourth grade in 2005.<sup>2</sup>

Retaining students has important consequences both for the individual as well as for schools. Overall, every repeater has the same effect on school resources as enrolling an additional student at that grade and subsequent grades and either leads to compromising per pupil school inputs e.g. through larger class size or to a pressure on public finances through the additional demand for teachers, classrooms, desks and other inputs.<sup>3</sup>

Opponents of grade repetition contend that it negatively impacts the retained individual by stigmatizing them and harming their self-esteem, by impairing established peer relationships and generally alienating the individual from school, which may in turn negatively affect academic achievement and increase the probability of dropping-out of school (Holmes 1989). Furthermore, repeating grades delays entrance of students

---

<sup>1</sup> 40 out of 43 African countries for which data is available in 2006 use grade retention (and for which average repetition rates exceed 4% in primary school) and 18 out of 23 Latin American and Caribbean countries.

<sup>2</sup> Data available at <http://stats.uis.unesco.org/unesco/ReportFolders/ReportFolders.aspx>. UNESCO Institute for Statistics, Data Centre, January 2012.

<sup>3</sup> A very rough estimate of the annual cost of repetition on public finances in Brazil using average expenditure per pupil at primary schools in 2006 of \$554 (in constant 2005 US\$) and 18,661,000 students enrolled at primary school and an average repetition rate over all grades of 18.7% (not accounting for loss of students due to drop-out etc.) amounts to approximately 1.9 billion US\$ (all data from UNESCO 2008).

into the labour market which poses substantial monetary cost on students over the life-cycle. In contrast, proponents argue that repetition can improve academic achievement by exposing low performing students to additional teaching and by allowing them to catch up on the curriculum and the content of teaching. This is particularly important if school absence for reasons such as illness in a given school year is the reason for retention. Grade retention may also help to make classes more homogeneous in achievement and therefore easier to teach by improving the match between peers in the classroom (Manacorda 2012).

There is a small but growing literature on estimating the causal effect of retention on subsequent educational outcomes (Gomes-Neto and Hanushek 1994, Eide and Showalter 2001, Dong 2009, Jacob and Lefgren 2004 and 2009, Manacorda 2012 and Glick and Sahn 2010). The results are mixed, with positive as well as negative estimates of the effect of repetition on academic achievement and school drop-out, and the results seem to depend critically on context and age of students.

Considering these mixed empirical findings on the effect on repeaters, the use of public resources and the undesirable consequences for public finances, the persistence of grade retention regimes in many countries is puzzling. This is particularly the case for developing countries where repetition rates are often very high and pressure on public resources is large. Furthermore, repetition increases the age variation in the classroom and repeaters may also directly lead to negative externalities on their peer students (Manski 1993, Lavy, Paserman and Schlosser 2012).

A possible explanation for the persistence of grade retention in many countries may be based on the deterrence effect of grade retention.<sup>4</sup> Grade retention induces students to exert effort as it potentially inflicts substantial costs of repetition on low

---

<sup>4</sup> Manacorda (2012) is the first to point out such a deterrence effect of retention in the literature. A related argument of a deterrence effect is discussed by Angrist et al. (2002) in relation to school vouchers and by Jacob (2005) in relation to high stakes testing in the US.



performers. The ex-ante threat of retention may therefore incentivize students to study in order to avoid being retained. This incentive effect of grade retention may have an important effect on mean student outcomes, as it is not restricted to repeaters only, but may create incentives for a much wider range of students. While the empirical literature on grade retention focuses on the ex-post effect on repeaters, there exists – to the author’s knowledge – no research on the ex-ante effect of the promotion regime on academic outcomes of a wider set of students. This analysis examines the effect of removing the deterrence of retention rather than estimating the effect of repetition on repeaters. Automatic grade promotion has been introduced in Brazil on a large scale since the early 2000’s partly to accelerate progress towards meeting the Millennium Development Goal of universal primary education and to reducing the cost of larger student cohorts (UNESCO 2012). I exploit credible exogenous variation in the timing of the adoption of automatic promotion for identification in a difference-in-differences (DiD) setting.

I find that the introduction of automatic promotion significantly reduces academic achievement measured by math test scores of fourth graders by 6.7% of a standard deviation. Quantile DiD results show that the strongest treatment effect can be found for the lower part of the test score distribution with considerably smaller effects in the tails of the distribution. This is consistent with an interpretation of the estimates as a disincentive effect of automatic promotion and the paper provides additional evidence in support of this interpretation. There is no evidence that the results are caused by teacher responses to the introduction of automatic promotion. Teachers are no more or less likely to assign and correct their students’ homework, and class size is unaffected by the policy introduction. Because there is only limited information on teaching practices available it is not possible to rule out completely the possibility of systematic teacher responses to the policy. The timing of the policy change limits the potential for

changes in the student composition of the test cohorts and I provide strong evidence that the socio-economic composition is unaffected by the policy and unlikely biases the estimates. There is also no evidence that the estimates are affected by systematic changes in student mobility across schools or by strategic test taking behaviour.

The remainder of the chapter is organized as follows. Section 1.2 provides information on the school system in Brazil and in the state of Minas Gerais. Section 1.3 presents the data. Section 1.4 describes the natural experiment and outlines the assignment of schools to treatment. Section 1.5 introduces the empirical strategy. The results, their interpretation and falsification exercises are presented in section 1.6, and section 1.7 concludes.

## **1.2 The school system in Brazil and Minas Gerais**

Primary school is compulsory in Brazil for children between the ages of 7 to 14 and consists of eight years of schooling (MEC 1996).<sup>5</sup> Public schooling is free at all ages and enrolment in primary and secondary school is open to students of all ages.

The Brazilian educational system has undergone substantial changes during the last two decades and has achieved considerable progress in expanding access to education. Starting from a primary school net enrolment rate of 85% in 1991, Brazil achieves today almost universal primary school enrolment with a net rate of 95% (UNESCO 2008). Primary school completion and youth literacy rates have improved notably, but the country continues to suffer from high repetition and drop-out rates.<sup>6</sup>

The national conditional cash transfer programme Bolsa Família, formerly Bolsa Escola, which is a means-tested monthly cash transfer to poor households conditional on school enrolment and regular attendance among other conditions, plays a significant

---

<sup>5</sup> The school entry age has been lowered recently to 6 years and primary school has been extended to 9 years.

<sup>6</sup> The overall repetition rate in primary schools in Brazil in 2006 was 18.7% and the total drop-out rate for primary school 19.5% (UNESCO 2008).

role for the rise in school enrolment and attendance of school age children (de Janvry, Finan and Sadoulet 2006).<sup>7</sup>

This analysis focuses on the state of Minas Gerais, the second most populous state in Brazil with an estimated population of about 19 million (IBGE 2007). Minas Gerais contributes 10% to the Brazilian GDP and is among the most developed states in Brazil (OECD 2005). The education system of Minas Gerais is among the most advanced and in national performance tests students regularly perform among the top (INEP 2007).

According to state legislation, the State Secretariat of Education (SEE) has extensive authority to plan, direct, execute, control and evaluate all educational activities in Minas Gerais. Based on the far-reaching decentralization of education in Brazil, the SEE transfers authority to a large extent to Regional Authorities for Education (Superintendências Regionais de Ensino: SREs) and directly to the municipalities. SREs and municipalities therefore play a major role in the provision of schooling and the implementation of educational policies.<sup>8</sup> Municipal schools account for more than half (56%) of all primary schools and state schools, that are directly under the control of the SEE, account for 22% of all schools. Besides the public provision of education private schools play an important role and account for the remaining 22%.<sup>9</sup>

### 1.3 Data description

This study uses data from two sources. Information on school characteristics comes from the annual Brazilian school census that is conducted by the National

---

<sup>7</sup> The conditionalities of Bolsa Família require a minimum school attendance of 85% and extend to the fulfilment of basic health care requirements such as vaccinations of the children and pre and postnatal medical consultations for pregnant women. Monthly per capita income in the household cannot exceed R\$120 (US\$57 in 2006) to remain eligible for the transfer. See Lindert et al. 2007 for a comprehensive description of the programme.

<sup>8</sup> The installation of FUNDEF, a federal fund established in 1996 with the aim of redistributing state and municipal resources back to (mainly) municipalities according to student numbers contributed to the improvement of the control of municipalities over educational decisions. See de Mello & Hoppe (2005) for an analysis of FUNDEF.

<sup>9</sup> There are also 28 federal schools in Brazil which are under the direct control of the federal government; the single federal school in Minas Gerais has not been included in the dataset.

Institute for the Study and Research on Education (INEP) under the control of the Federal Ministry of Education (MEC). The Brazilian school census compiles data annually from all primary and secondary schools in Brazil. The exceptionally rich data includes information on the location and administrative dependence of schools, physical characteristics (quantity of premises and class rooms, equipment and teaching material), the participation in national, state and municipal school programmes, the number of teachers and administrative staff, average class-size, detailed information on student flows (number of students in each grade by to age, repetition, drop-out and student transfer rates) among other information.<sup>10</sup>

The school census also contains the information on the grade promotion regime of adopted in each school (grade retention versus automatic promotion), which is used to establish treatment and control groups.

The second part of the data comes from the State System of the Evaluation of Public Education (Sistema Mineiro de Avaliação da Educação Pública: SIMAVE), which includes the programme for the evaluation of state primary and secondary schools (Programa de Avaliação da Educação Básica: PROEB).<sup>11</sup> Results from the programme are used for the evaluation and design of educational policies in the state; the results are however not used by the schools to evaluate individual student performance, for example for the grade promotion decision.

The main outcome variable is student achievement in state schools in Minas Gerais measured by math test scores in 2003 and 2006. All classes and all students in fourth grade of each school are examined and participation of schools and students is compulsory. The cognitive test scores are standardized to a mean of 500 and a standard deviation of 100. In total 246,959 students have been tested in 1,993 state schools in

---

<sup>10</sup> Summary statistics for the public schools used in this analysis are presented in panel A of table 9.

<sup>11</sup> Schools under the administration of the municipality or the federal government are not included in SIMAVE.

Minas Gerais. I use the repeated cross-section of test score data from 2003 and 2006 for this analysis. The students in the dataset have, as generally students in public schools, a deprived socioeconomic background. Almost half (45.6%) of the families with children at state schools in Minas Gerais qualify for Bolsa Família and can be considered poor. Information on sex, date of birth, racial background and on the socio-economic family background is also available from an adjunct questionnaire. Unfortunately, only the 2003 wave of the socio-economic questionnaire contains information on parental education. Panel B of table 9 presents summary statistics on these variables.

## **1.4 The General Education Act of 1996: the case of a quasi-experiment**

### **1.4.1 Policy background**

The General Education Act of 1996 (Lei de Diretrizes e Bases da Educação Nacional: LDB) paved the way for the introduction of automatic promotion policies in Brazil. Federal Law No 9.394/1996, which came into effect in 1998, regulates the responsibility for education between the federal, state and municipal level and facilitated federal and state programmes to control the grade promotion regime (Pino and Koslinski 1999). Section 3 of Art. 32 §1&2 formally distinguishes two alternatives for educational authorities to organize student progression: besides the conventional annual grade promotion regime the option of automatic promotion was introduced, a system in which students progress automatically to the next grade at the end of the school year. Between these two extremes, a mixture of both regimes was also permitted. In the mixed regime, schools define “learning cycles” that stretch over several – most commonly three – school years. During the initial years of the cycle students are promoted automatically. In the final year of a cycle students that fail to meet the minimum requirements set in the curriculum are retained. The idea behind learning cycles is to allow students an individual studying pace (Mainardes 2010). If students fall

behind their classmates they have a longer period to catch up on the curriculum. This particularly aims at reducing the long-run impact of negative temporary shocks, such as school days lost to sickness or adverse family events. For mixed regime schools that have adopted automatic promotion in learning cycles, grade retention is not entirely eliminated, but limited to the final year of the cycles. The LDB furthermore sets fundamental criteria on how to organize promotion under any one regime: In every school year a minimum attendance of 75% of all school days must be fulfilled as a general requirement for promotion, so that grade retention is still permitted if students fall below a 75% minimum attendance.

According to the legal framework of the LDB the decision on the promotion regime and its exact specifications is taken on the state level. Automatic promotion was introduced at an early stage by the states of São Paulo, Minas Gerais and Paraná, to some extent in the state of Pernambuco and by the federal resolution SE No 4, 1/98 in all federal schools in Brazil. A recent federal resolution disallowed retention for the first three schools years in all schools in Brazil from 2011.

In the state of Minas Gerais the new regime has been established by state resolution No. 8.086 in 1997. It stresses the autonomy of each public school in the decision whether to continue with the annual repetition regime or to introduce automatic promotion. In the year 2000 1,449 out of 1,993 state schools had established automatic promotion with two initial three-year cycles. At the beginning of the school year 2004 the remaining 544 state schools switched to automatic promotion.

#### **1.4.2 Assignment to treatment**

Schools that adopted automatic promotion at the beginning of the year 2004 make up the treatment group and schools with automatic promotion (which have adopted automatic promotion since the year 2000) the control group. I focus on two cohorts of

fourth graders, which I call the test cohorts 2003 and 2006 for which test scores are available. Figure 2 presents an overview of these two cohorts and the change in the organization of promotion for the control and treatment group.

When using this division into treatment and control group for comparison a sound understanding of the assignment process that leads to this division is essential. In the case of state schools in Minas Gerais the 46 regional authorities for education needed to propose a plan of implementation of automatic promotion for the schools under their administration. The decision for early adoption of the policy was made by each SRE in agreement with the state secretariat. The second wave was initiated by the SEE in an attempt to introduce automatic promotion universally in all schools. As the adoption of the policy is not randomized across schools in an experimental setting, treatment and control schools may not be balanced in the distribution of school and student characteristics. Although the identification strategy used in this analysis does not rely on the distribution of covariates being balanced, it is generally reassuring to find school and mean student characteristics of treatment and control group to be very similar. Table 9, panel A and B present descriptive statistics of treatment and controls schools for 2003 and 2006. T-tests (and Chi-square for categorical variable) for the equality of means between treatment and comparison group, accounting for clustering on the SRE level, reveal only very few small but statistically significant differences. As sample size is partly reflected in the t-statistics, it is more useful to look at the normalized difference  $norm - diff = \frac{\bar{X}_0 - \bar{X}_1}{\sqrt{S_{X,0}^2 + S_{X,1}^2}}$  between means by treatment status as a scale-free measure of the balancing properties of the covariates (Imbens and Wooldridge 2009). The normalized difference is small for all covariates and never exceeds the absolute value 0.25,<sup>12</sup> suggesting that treatment and control schools are indeed very similar in terms of

---

<sup>12</sup> This is a rule of thumb suggested in Imbens & Wooldridge 2009 to check the unconfoundedness assumption for the use of linear regression in estimating average treatment effects.

their physical school characteristics. Even more importantly, the normalized differences of mean student characteristics, which may indicate compositional differences of the student populations, are all very small and are far below the suggested rule-of-thumb value of  $|0.25|$  in both years. Apart from mean age, which differs slightly as expected,<sup>13</sup> no other variable reveals any considerable difference at the mean. The overlap in the covariate distributions can also be examined by looking at the distribution of the propensity score for the treatment and control group. Figure 1 shows the propensity score for the probability of treatment for the treatment and control group revealing substantial overlap in the multivariate distribution of covariates and a relatively similar pattern of the distribution of the propensity score for the treatment and control group.<sup>14</sup>

In addition, I estimate a linear probability model to determine whether there are systematic differences between schools that have adopted automatic promotion at different points in time and to learn what observable school characteristics – if any – determine early adoption. The results are presented in table 14. The coefficients on the set of school characteristics are generally small and only very few are statistically significant. When including SRE controls even fewer variables show a significant effect and it is difficult to establish any systematic pattern.

Given the similarity of treatment and control schools with respect to school characteristics and the student composition, it is plausible to consider the assignment of schools to treatment and control groups as conditionally random.

### 1.5 Empirical strategy

To estimate the treatment effect of the policy change I use a DiD estimator exploiting the variation in treatment status of schools over time, identifying an average

---

<sup>13</sup> Mean age is expected to differ as part of the treatment, which will be clarified in section 6.3.

<sup>14</sup> A formal test under the null for the equality of the distribution (Kolmogorov-Smirnov) of the propensity score is nevertheless rejected.



treatment effect on individuals in schools assigned to treatment. The double difference approach is capable of removing biases resulting from permanent latent differences between treatment and control as well as biases resulting from common trends over time. The estimation in a regression setup allows including additional regressors on the individual and school level to improve precision and to test for the presence of omitted-school specific trends, in particular related to potential changes in the student composition. Identification requires that trends in student outcomes at treated and control schools would not be systematically different in the absence of treatment.

Under this identifying assumption, I estimate the effect of the introduction of automatic promotion on test scores of fourth graders by the following regression model:

$$Y_{ist} = \beta_0 + d_s + d_t + \gamma d_{st} + \beta_1 Z_{it} + \beta_2 X_{st} + \epsilon_{ist} \quad (1)$$

where  $Y_{ist}$  is the test score for individual  $i$  in school  $s$  at time  $t$ ,  $d_s$  is a school dummy which captures school-specific time invariant effects,  $d_t$  is a time dummy which captures the common time trend of control and treatment group,  $d_{st}$  is the time/treatment-status interaction term containing information on the treatment status of schools, that varies over time.  $\gamma$  in equation (1) is the coefficient of interest and reflects the average treatment effect of the introduction of automatic promotion on test scores of fourth graders.  $Z_{it}$  is a set of covariates controlling for individual characteristics.  $X_{st}$  denotes a set of exogenous covariates for class and school characteristics, including average socioeconomic characteristics of students, detailed school characteristics,<sup>15</sup> the participation in federal, state and municipal educational programmes,<sup>16</sup> teacher

---

<sup>15</sup> Specifically, the covariates include initial (first grade) enrolment, number of teachers at school, number of total staff (besides teaching staff), dummy variables describing the type of the premises used for the school, dummies for the availability and number of teaching material (e.g. overhead projectors, personal computers, TV and video sets etc.), the availability of computer and science labs, school kitchen, the quality of sanitary units, number of class rooms in- and outside the school and dummies for whether the school provides all 8 years of primary education.

<sup>16</sup> These programmes include National Minimum Income Programme, Free School Lunch programme, the provision of public school transportation, TV escola (a national education TV programme), other

characteristics and other.<sup>17</sup>  $\varepsilon$  is a stochastic error term. Although non-random assignment of schools to treatment may lead to a correlation between assignment status and outcomes, this does not violate the common trend assumption as long as any differences that lead to the adoption of the policy are captured by the school-fixed effects. The common trend assumption may nevertheless be violated if selection into treatment was based on pre-treatment trends in school characteristics that differ between treatment and control. If, for example, schools with high performing students and low repetition rates adopt automatic promotion test scores and treatment status are correlated for reasons other than the treatment impact of automatic promotion. Unfortunately I do not have pre-treatment test score data to test directly for the common trend assumption. I nevertheless can investigate whether selection into treatment is based on pre-treatment differences in repetition rates. Table 10 reveals that pre-intervention repetition rates (from the 1997 school census before automatic promotion was introduced at any school) were virtually identical across treatment and control schools, so that there is little concern for self-selection of schools into treatment based on high or low repetition rates. The table also reports pre-treatment class size (averaged over grades 1-4) and pre-treatment student-teacher ratio (averaged over grades 1-4). While there is small difference in class size of about one student, there is virtually no difference in the student teacher ratio and the normalized difference for both variables is well below  $|\cdot 25|$ . Classroom capacity constraints therefore were unlikely the driving factor behind the decision for early adoption.<sup>18</sup> As I have pointed out earlier, the first wave of the policy adoption was initiated on the SRE level, which furthermore limits the potential for

---

educational TV programmes, computer literacy programmes, and other state and municipal school programmes.

<sup>17</sup> This will also allow accounting for eligibility specific effects (Ashenfelter 1978). This way the above time invariant composition assumption can be relaxed to accommodate for the case where treatment and control group are expected to differ in covariates that may affect the outcome variable.

<sup>18</sup> Lam and Marteleto (2006) show that the demographic transition in Brazil in the 1990's had a strong impact on student cohort sizes and enrolment rates in Brazil, but this does not seem to be relevant in the context of this study.

individual schools to select into treatment based on trends in test scores. The second wave was then determined by the decision of the CEE made for all remaining schools, so that there is virtually no scope for selection on a pre-treatment trend basis.

As the treatment regressor varies at the school level and test scores of students in the same school are likely correlated, for example because they share the same learning environment and/or are from the same neighbourhood, conventional standard errors may be misleading as they do not account for the grouped error structure. The robust standard errors reported therefore allow for clustering on the school level (Donald and Lang 2007).

## **1.6 Estimation results**

### **1.6.1 Main results**

The basic idea of the DiD strategy can be illustrated by a simple 2-by-2 table. Table 1 shows the levels and differences in test scores between treatment and control groups and the changes over time. The first row reports means before treatment (year=2003), when control schools were already under the automatic promotion regime and the treatment schools were still under the annual grade retention regime and the mean difference for the two groups. The entries in the first column reveal that schools that had already adopted automatic promotion have a mean score that is 7.05% of a standard deviation lower than schools that had not yet adopted the new regime in 2003. After the adoption of automatic promotion by schools of the treatment group this difference almost completely disappears and students at both groups have very similar average test scores and the difference in means is not statistically significant. Likewise, schools in the control group have very similar mean test scores over time with a difference that is not significantly different from zero. The lower right entry reports the simple DiD estimate, which can be interpreted as the causal effect of treatment under

the above identifying assumptions. The adoption of automatic promotion leads to a decrease in test scores of 6.65% of a standard deviation. Almost the entire fraction of the DiD outcome originates from the pre-treatment difference between control and treatment schools. After the adoption of automatic promotion in treatment schools the difference between treated and control schools almost completely disappears.

This simple double difference can be amended in a regression framework following equation (1) to improve precision of the estimates and to be able to control for covariates and check the sensitivity of the estimates to their inclusion. Table 2 presents the estimates for different sets of controls. All specifications include school fixed effects and year dummies. School fixed effects capture stable unobserved characteristics of the schools and year dummies control for common trends in the test scores that are not related to treatment. Specification (1) of table 2 includes school controls, specification (2) controls for school and peer characteristics and specification (3) also includes individual level covariates. The estimates in all specifications reveal a stable negative effect of around 6% of a standard deviation and are very precisely estimated (1% level of significance). Adding school level and peer controls reduces the negative effect, but the reduction is relatively small. Controlling additionally for individual characteristics delivers estimates of virtually the same size as the simple double difference in table 1. The results reveal that the regime change from annual grade retention to automatic promotion has a significant negative impact on educational attainment on fourth graders in state schools in Minas Gerais. In the next section I will discuss the interpretation of the results.

### 1.6.2 Interpretation of the results and the disincentive effect of automatic promotion

Table 3, column 1 reports the DiD estimates of the treatment on repetition rates for grades 1-4, following the tested cohorts of 2003 and 2006 over grades 1-4. The bottom entry for column 1 shows how the introduction of automatic promotion reduces the repetition rate in fourth grade by 0.086. Prior to the policy change, about 10% of all students in treatment schools repeated fourth grade, but only about 2% did so after the introduction of automatic promotion.<sup>19</sup> In this analysis I am interested in understanding whether the estimated effect on test scores can be explained by the elimination of the threat of retention for fourth grade students. The two cohorts of students at treatment schools face indeed very different incentives from the grade promotion regime; while the 2003 cohort is subject to grade retention, the 2006 cohort does not face the threat of being retained.

If the estimated effect is caused by a change in study incentives to avoid being retained, one would expect heterogeneous treatment effects along the test score distribution. Students in the lower tail of the distribution should be more heavily impacted by removing this incentive when compared to students in the upper tail, as these students should be less concerned about the possibility of retention. For that purpose I estimate equation (1) applying DiD to each quantile instead of the mean under analogue assumptions to the standard DiD (Koenker 2004, Athey and Imbens 2006, Firpo et al. 2009).<sup>20</sup> Table 4 provides the quantile DiD estimates and reveals substantial differences in the treatment effect across the nine quantiles. The estimates range between -9.01 and -3.92 and are more pronounced in the lower half of the distribution, with the strongest effects centred on the fourth quantile. The effect of automatic

---

<sup>19</sup> Repetition rates stay above zero because repetition is still possible when failing to achieve 75% minimum school attendance.

<sup>20</sup> Recent applications of quantile panel methods include Havnes and Mogstad 2010, Gamper-Rabindran, Khan and Timmins 2010 and Lamarche 2011.

promotion is much smaller for the top two quantiles and not statistically significant, yet still negative and non-negligible in magnitude. The inverse u-shaped distribution of effects is consistent with the interpretation of the estimates as disincentive effect of automatic grade promotion, such that the treatment effect is largest for students left of the centre of the distribution close to the assumed grade promotion threshold and smaller for high performing students that are unlikely to be retained. Similarly, for students at the very bottom of the distribution the effects are somewhat smaller with a coefficient of -7.49 but still above the mean treatment effect. The slightly smaller effect at the bottom of the distribution could be explained by either a different perception of the cost associated with retention or the fact that grade retention is a possibility for these low performing students regardless of their effort.<sup>21</sup> There is some suggestive evidence that automatic promotion indeed directly impacts the behaviour of students and reduced their study effort. Column (1) of table 11 reports the effect of the policy introduction on the propensity of students doing their homework.<sup>22</sup> The DiD estimates show that after the introduction of automatic promotion fewer of the children do their homework (a decrease of 0.014); the coefficient is only marginally significant though. Interestingly, the change in the retention regime also changes the parents' involvement with their children's homework. Column (2) of the table shows that parents are more likely to help with their children's homework (an increase of 0.022). This reveals that parents may well be aware of the disincentive from automatic promotion and they may try to counteract the potential reduction in their children's study effort. If anything, increased parental involvement would however bias the estimates in table 2 towards zero, rather than explain the estimated effect.

---

<sup>21</sup> Separate estimates by socio-economic status as proxied by the number of books in the household do not reveal heterogeneous effects along that margin (results not reported).

<sup>22</sup> All of the outcome variables in table 11 are based on pupil reported behavioural responses of themselves, their parents and their teachers and should therefore be considered more cautiously.

The distribution of treatment effects in table 4 is also consistent with an explanation based on changes in teacher incentives from automatic promotion. Teachers may for example focus less on students in the bottom half of the distribution if they previously cared about them being promoted.<sup>23</sup> Information on whether teachers assign and correct homework may shed some light on potential teacher responses to automatic promotion. Columns (3) and (4) of table 11 report the DiD estimates for teachers assigning and correcting homework respectively. Both coefficients are very close to zero and not statistically significant so that there is no evidence that teachers respond systematically to automatic grade promotion.<sup>24</sup>

### **1.6.3.1 Changes in the student composition**

For the interpretation of the estimates as disincentive effect, any possible channel of effect of automatic promotion on outcomes – other than the disincentive effect – has to be precluded. Most importantly, potential changes in the composition of students in treatment and control schools over time could systematically lower test scores rather than the changes in incentives reducing the effort of students. Because there is grade retention in control and treatment schools in both periods at the end of third grade this leads to a positive selection of the students entering into fourth grade in both treatment and control schools and this mechanically limits the potential for changes in the student composition. In the section 6.4 I will discuss the implications of this in detail.

Table 6 reports DiD estimates for a range of mean socioeconomic variables; for each outcome variable I have fitted a separate regression including school fixed effects and year dummies. Only the coefficients on the mean number of fridges per household and on mean age are statistically significant. All other indicators of the socio-economic composition are not affected by the introduction of automatic promotion, which is very

---

<sup>23</sup> Teachers may equally worry about the lost incentive for students and target their effort on the most affected students, so that a potential teacher response may go either way.

<sup>24</sup> Section 6.8 looks separately at class size as another teaching input.

reassuring. While the coefficient for the mean number of fridges per household is very small and may be due to some spurious correlation, the significant reduction in mean age by about one month is more relevant and it is important to understand the source of this reduction in age and its consequence for the interpretation of the result.

This reduction in age is caused by the difference in the inflow of repeaters in fourth grade at the treatment schools before and after treatment. Whereas treatment schools still received an inflow of repeaters from fourth grade of the previous year at the beginning of the year 2003, there was no such inflow of repeaters in 2006, which leads to the reduction in mean age, as repeaters are on average one year older. Table 8, column 2 shows the DiD estimate of the policy change on the net inflow of students from first to fourth grade and from first to third grade in column (1). Whereas the coefficient in column (1) is very small, negative and not statistically significant, the coefficient for the net inflow of students including the inflow of repeaters from the previous year at the beginning of fourth grade is sizeable, positive and very precisely estimated (column (2)). Looking at the direct effect of the inflow of repeaters on mean age of the cohort reveals that this almost exactly explains the age effect estimated in table 6.<sup>25</sup> This means that the composition is altered due to treatment and it is important to understand the potential bias of the compositional change on mean achievement.

Even assuming a positive effect of repetition on educational outcomes of repeaters,<sup>26</sup> it is very plausible to assume that average performance of repeaters is still below the mean performance of non-repeaters in the test cohort, as repeaters are selected as the lowest performers in fourth grade in the preceding year.<sup>27</sup> How does this differential inflow affect the outcome variable of interest? As there was an inflow of such low performing students in 2003, but not in 2006 the results for the disincentive effect of automatic

---

<sup>25</sup> Assuming that they are about one year older the inflow of repeaters at fourth grade leads to a decrease of mean age of the cohort of 36 days compared to the estimated effect on mean age of 39 days.

<sup>26</sup> And a direct effect related to age, as repeaters are one year due to repeating the grade.

<sup>27</sup> This is confirmed by the findings elsewhere; see Manacorda (2012) for example.



promotion are, if anything, biased towards zero and the reported coefficients in table 2 (panel A) need to be regarded as a lower bound of the true effect. Unfortunately, there is no direct information in the student questionnaire on whether and when students were retained. I can nevertheless use individual age to single out repeaters to some extent. A regression sensitivity analysis that includes individual age as a control variable may give an idea about the size of the bias from the differential inflow of repeaters. Adding individual age to specification (1) leads to an increase in the negative effect of about 20% to -7.97% of a standard deviation compared to -6.65 % without controlling for age, reported in panel B of table 2. Controlling for individual age in specification (2) and (3) leads to a very similar increase of 20% of the effect to -7.33% and -6.77%, respectively. An alternative way of investigating the importance of the bias for all specifications is to restrict the estimations to students that have never repeated by excluding all students outside the target age range of fourth graders. Once students from the additional inflow at fourth grade from the sample are removed, this leaves a sample of students that have never repeated.<sup>28</sup> Panel A of table 5 reports the results for the same specifications as in table 2, but restricts the sample to students in the target age range for fourth graders. By restricting the sample in this way the coefficients exceed the estimates of the original full sample in all specifications by around 30%. The estimated effect is a further 11-16% larger compared to the estimates in panel B of table 2. Restricting the sample to repeaters (panel B, table 5) reveals a negative effect that is considerably smaller and no longer statistically significant. The number of excluded students is nevertheless larger than what could be explained by excluding fourth grade repeaters only, as removing overage students from the sample also removes students that have repeated at third grade. As repetition is equally possible in all schools at third grade, the additional

---

<sup>28</sup> Nevertheless I cannot distinguish repeaters from students that have enrolled late at first grade. With rather strict enforcement of the enrolment age in Minas Gerais and the incentives to parents to enrol their children based on Bolsa Família conditions, late enrolment is nevertheless rather limited.

increase in the estimates is therefore not necessarily related to treatment. The increase rather suggests that the incentive of grade retention may have a different impact on previous repeaters compared to students that have never repeated a grade. The cost of repetition is likely highest for students that have not previously repeated. In contrast the cost of being retained again is smaller for previous repeaters, as they may already have suffered stigmatization and have already been separated from their original peer group. The difference in results for the restricted sample therefore may not only reflect the correction for the differential inflow of repeaters at fourth grade, but may also more generally reflect heterogeneous effects on repeaters and non-repeaters.

A more comprehensive analysis of the sensitivity of the estimates to the inclusion of age controls is provided in table 12. I present different specifications of equation (1) with and without controlling for individual age for the full sample (panel A) and the age restricted sample in panel B. The results support the previous findings. Adding individual age as control (columns 4, 6 and 8) strengthens the negative effect in the full and restricted sample for all the different specifications.

Besides the direct effect on the composition, there may be an indirect effect on students of having repeaters in the class room. Repeaters may impose a negative externality on their peers because their achievement is lower or because they may be more disruptive in class (Lazear 2001). Lavy, Paserman and Schlosser (2012) elaborate on the extent of ability peer effects associated with repeaters and show that academic performance and behaviour of repeaters may be responsible for the negative effect. By adding peer age in the DiD specification I can control for potential peer effects from the differential inflow at fourth grade. Adding peer age as control only moderately increases in size the coefficients in specification 1 and 2 in panel C of table 2. Columns 2 and 5 in panels A and B of table 12 reveal that the inclusion of peer age only has a minor effect when controlling for other peer variables and does not strengthen the estimates of the

treatment effect suggesting that there is no noteworthy bias on the estimates. If anything, a negative peer effect of repeaters, as suggested in the literature, would lead to an underestimation of the effect.

Conditioning on individual age and restricting the sample to non-repeaters reveals that the differential inflow at fourth grade changes the composition of students in a way that underestimates the true effect of automatic promotion by not taking into account the net inflow of repeaters into fourth grade in 2003. The size of the downward bias ranges between 20% and 30%. The estimates for the restricted sample should nevertheless be considered with caution, as the disincentive of automatic promotion may have differential impact on previous repeaters and non-repeaters.

#### **1.6.3.2 Introduction of automatic promotion at 2<sup>nd</sup> grade**

Because of the introduction of automatic promotion in treatment schools at second grade of the cohort of interest, this potentially may also have an impact on the composition of students. Table 3 reveals how repetition rates from first to fourth grade of the theoretic test cohorts are affected by the policy introduction. The estimates for first and third grade show no effect of the treatment as expected. Rates at first grade are unaffected with the policy change occurring only in the subsequent year and rates in third grade are unaffected as the final year of the cycle remains with grade retention for both cohorts in treatment and control group. The estimate for the impact on the second grade reveals how the policy introduction lowers the repetition rate by almost 12% at second grade in 2004. The potential threat to the interpretation of the results arises from the fact that by introducing automatic promotion at second grade for the 2006 exam cohort, this cohort may be “contaminated” by low performers that would have been removed in the absence of treatment. The mean repetition rate for second grade at treatment schools drops from 12.8% (2003 exam cohort) to 3.1% (treatment cohort).

Rather than looking at second grade repetition only, in- and outflows in each grade up to the end of third grade have to be considered when examining the effect of the adoption of automatic promotion on the student composition. Looking at overall student flows of the test cohorts reveals that the negative selection has largely cancelled out when the test cohort enters fourth grade in 2006. In particular, repetition in third grade plays an important role here. The first column of table 8 reports the effect of the policy introduction on the net flow taking into account in- and outflows over grades 1-3. The net inflow due to the introduction of automatic promotion is very close to zero and not statistically significant. This is mainly based on two factors: Focussing only on treatment schools, table 13 shows that repetition rates at third grade actually increased by about 4.3% for the 2006 exam cohort, which filters out a substantive fraction of the low-performers already. Furthermore, third grade repetition rates for the two cohorts have to be compared with caution, as these may have a different impact on removing low-performers from the previous year depending on the inflow of students into third grade at the beginning of the year. Considering net-flows, the composition of 2003 and 2006 cohorts are practically unaffected at the beginning of fourth grade. As mentioned earlier, the socioeconomic composition between the cohorts (table 6) is virtually unaltered by treatment, which supports the premise, that the policy introduction does not change the composition of students up to fourth grade.

This is also corroborated by the fact that almost the entire fraction of the DiD result arises from the ex-ante difference between treatment and control group in 2003, rather than from the difference after treatment. The results for the simple difference over time of the control schools and the difference between control and treatment schools after treatment in 2006 are very small and not significant at conventional levels.

### 1.6.3.3 Effect of the policy change on drop-out rates

The effect of retention on student drop-out has been studied elsewhere in the literature (see Jacob and Lefgren 2004, 2009, Manacorda 2012). If the introduction of automatic promotion has an effect on drop-out rates in grades prior to fourth grade, this may change unobserved student characteristics that cannot be controlled for. I estimate the effect of the introduction of automatic promotion on drop-out rates in a DiD specification similar to equation (1) as  $Y_{st} = \beta_0 + d_s + d_t + \gamma d_{st} + \epsilon_{st}$  (2), using aggregated data from the school census. Column (2) of table 3 reports the coefficients for each grade. Drop-out rates in second grade are unaffected by the policy change.<sup>29</sup> The treatment nevertheless has a small effect on drop-out rates at third grade, by reducing the rate by half a percent. This is equivalent to a mean reduction of 0.31 students per school/cohort and presumably negligible in its potential impact on mean test scores.

### 1.6.3.4 Effect of the policy change on school transfer rates

Another potential source for compositional changes is related to student mobility between schools. Parents that expect a negative effect of automatic promotion on their children may for instance want to move their children to a school with grade retention. In Minas Gerais the possibility for switching public schools is limited as enrolment is mainly based on residence and a single public school often serves the local neighbourhood. Given very substantial fees at private schools it is also unlikely that parents move their children into private schools to avoid a specific grade promotion regime. As the policy was introduced while the cohort of interest was in second grade the incentive for parents to move their children is further reduced. To test for any effect

---

<sup>29</sup> First grade repetition rates are also unaffected as predicted, because the policy change only takes effect after first grade. This is a relevant observation as it shows that there are no anticipatory effects from schools to the introduction to the policy change.

of the policy change on between-school mobility I estimate the effect of the introduction of automatic promotion on student transfer rates using the same framework as in the previous section. Columns (3) and (4) of table 3 report point estimates for outgoing and incoming transfer rates that are close to zero and not significant for any grades, so that there is no evidence that student mobility has an impact on the student composition.

#### **1.6.3.5 Systematic test taking behaviour**

Although participation in PROEB is mandatory on the school and individual level, some students fail to attend the test.<sup>30</sup> If the propensity to show up at the exam is related to the capacity of the student and to the treatment status of the school, this may bias the estimates. This might be induced by strategic behaviour of school administrators or teachers trying to manipulate the mean test scores of their school in the PROEB exam. If this is systematically linked to treatment status this could bias the estimates. Notably, individually identified test results are not available to the schools and PROEB test scores are not used by schools for the grade promotion decisions. I use information from the official student numbers in each school from the school census and compare these to the number of students participating in PROEB. I estimate equation (2) using the difference between the two figures as outcome variable. Table 7 presents the results from the regression. The coefficient is very small (0.119 students) and not statistically significant so that there is no evidence for systematic absence of students from the test.

#### **1.6.3.6 Effect of the policy change on class size**

There may be other teaching inputs that could be affected by the policy change; for example a reduction in retention rates may affect class-size, which in turn may have

---

<sup>30</sup> The participation rate for the 2003 and 2006 wave of PROEB is around 95% as participation is strictly enforced and absence is only permitted in case of illness.

an impact on outcomes. There is a comprehensive literature on the effect of class size on student performance but the overall picture about class-size effects remains rather unclear.<sup>31</sup> To rule out that the estimates are biased by an effect of the policy on class-size I test for changes in class-size for each grade induced by the policy change for the cohorts of interest and column (5) of table 3 reports the DiD results. There is no evidence for an effect of the policy change on class-size in any grade, so that estimates on test scores are unlikely biased by treatment induced class-size effects. Even under the assumption that the introduction of automatic promotion releases other school resources that could be allocated to fourth grade students (for which there is no evidence in the present analysis) this would lead to underestimating the true impact of the disincentive created by automatic promotion.

The fact that none of the above estimates (for repetition rates, drop-out rates, class-size, transfer rates) reveal any significant effect for first grade estimates is in itself an important falsification exercise. All these estimates are based on a placebo-treatment as the first grade of the 2006 exam cohort was not yet affected by the policy introduction. This also indicates that there are no anticipatory effects of the schools in respect to the imminent introduction of automatic promotion that may affect student outcomes at a later stage.

## 1.7 Conclusions

Existing empirical work on grade retention has to date focused on analysing the direct effect of retention on repeaters. The focus on the ex-post effects may nevertheless neglect an important effect of the grade retention regime that works through incentives to study on a larger range of students than repeaters only. The introduction of automatic promotion removes the incentives linked to the threat of retention and in this paper I use

---

<sup>31</sup> See Hoxby (2000) and Angrist & Lavy (1999) for two prominent studies on class size effects.

exogenous variation in the timing of the policy adoption in public primary schools in the Brazilian state of Minas Gerais to obtain causal estimates of automatic promotion.

Using a DiD approach I find a negative effect of about 6, 7% of a standard deviation, significant at the 1% level. Controlling for individual age strengthens the negative effect by about 20%, which gives an idea about the size of the bias associated with the differential inflow of repeaters into fourth grade before and after treatment. The estimated effect of the introduction of automatic promotion is of non-negligible size. Considering that automatic promotion may have a negative effect in several grades, the overall impact of the automatic promotion regime may lead to considerable loss of academic achievement over the eight years of primary school. Quantile DiD estimates yield an interesting insight into the distribution of effects. The quantile DiD estimates reveal that a large set of students is impacted by the policy change and not only the least performing students. The inverse u-shape of effects along the test score distribution is consistent with an interpretation of the estimates as disincentive effect of automatic promotion. Some further suggestive evidence on student responses support this interpretation. Other potential channels, in particular related to changes in the student composition, can be ruled out.

The estimation of a disincentive effect associated with automatic promotion closes a gap in the literature on the effects of grade retention and helps to explain the persistence of repetition regimes in many countries. Grade retention reduces internal flow efficiency at schools and is a costly policy, but may have a positive effect on academic achievement through a deterrence effect. Rather than focusing only on the effect on repeaters, attempts to assess the costs and benefits of grade retention therefore need to take into account the effects on non-repeaters as well.

This is important in the light of the universal introduction of automatic promotion in all primary schools in Brazil that came into effect by federal legislation in 2011.



Although the Brazilian experience may not be completely transferable to other countries – often with lower repetition rates – the findings may nevertheless be relevant for countries facing pressures to meet the Millennium Development Goal of universal primary education and who may regard the introduction of automatic promotion as a suitable way to reduce repetition rates and increase school completion rates.

**Table 1: Test Score Means in Treatment and Control Schools  
before and after the Adoption in the Treatment Schools**

	Before treatment	After treatment	Change in mean test scores
Control schools	498.48 (1.55)	498.99 (1.51)	-0.51 (2.01)
Treatment schools	505.53 (2.71)	499.39 (2.59)	6.14 (2.97)
Difference in mean test scores	-7.05 (3.12)	-0.40 (2.99)	-6.65 (3.22)

Notes: Mean outcomes for treatment and control before and after treatment. Standard errors, adjusted for clustering within SREs, are reported in parenthesis.

**Table 2: Main Estimation Results and Sensitivity to Age Controls**

Dependent variable: PROEB math test scores

Observations: 244,081, number of clusters: 1,993

	(1)	(2)	(3)
<b>Panel A</b>			
Treatment effect	-6.13*** (2.03)	-5.67*** (1.98)	-6.24*** (2.04)
R-squared	0.01	0.01	0.17
<b>Panel B – adding individual age control</b>			
Treatment effect	-7.33*** (2.11)	-6.77*** (2.04)	-6.24*** (2.04)
R-squared	0.04	0.04	0.17
<b>Panel C – adding peer age control</b>			
Treatment effect	-6.46*** (2.02)	-5.67*** (1.98)	-6.24*** (2.04)
R-squared	0.01	0.01	0.17
School fixed effects	yes	yes	yes
Year dummies	yes	yes	yes
School level controls	yes	yes	yes
Peer characteristics controls	no	yes	yes
Individual characteristics controls	no	no	yes

Notes: \*\*\* denotes significance at 1%. Robust standard errors, adjusted for clustering within schools, are reported in parenthesis. Specification (1) contains year dummies and school fixed effects, specification (2) additionally controls for a rich set of school characteristics (physical characteristics of the school and the class rooms, teaching material, teacher characteristics, participation in educational programmes etc.), specification (3) additionally controls for peer socio-economic characteristics at the school level and specification (4) also controls for individual characteristics.

**Table 3: Effect of the Adoption of Automatic Promotion on Student Flows and Class-size**

Dependent variable:					
	Repetition rate	Drop-out rate	Transfer-rate outgoing	Transfer-rate incoming	Class-size
	(1)	(2)	(3)	(4)	(5)
Grade 1	-0.010 (0.010)	0.001 (0.003)	-0.005 (0.007)	0.007 (0.008)	0.692 (0.363)
Grade 2	-0.118*** (0.010)	-0.002 (0.003)	0.002 (0.006)	0.003 (0.009)	0.828 (0.422)
Grade 3	-0.019 (0.011)	-0.005** (0.002)	-0.002 (0.003)	0.007 (0.009)	0.502 (0.416)
Grade 4	-0.086*** (0.007)	-0.005 (0.004)	-0.007 (0.004)	0.006 (0.008)	0.252 (0.336)

Number of schools: 1993, years 2000-2006, average cohort size: 61.24

Notes: \*\*\* denotes significance at 1%, \*\* denotes significance at 5%. The coefficients report the effect of introducing automatic promotion on the dependent variables for 1<sup>st</sup> to 4<sup>th</sup> grade using data from the school census 2000-2006 following the theoretical test cohorts. For each grade a separate regression has been fitted estimating the effect corresponding to equation (1) as  $y_{st} = \beta_0 + d_s + d_t + \gamma d_{st} + \varepsilon_{st}$ . The regression estimates are weighted by school cohort size and include year dummies ( $d_t$ ) and school fixed effects ( $d_s$ ). Robust standard errors, adjusted for clustering within 46 SREs, are reported in parenthesis.

**Table 4: Quantile Treatment Effects**

Dependent variable: PROEB test scores									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Quantile	1	2	3	4	5	6	7	8	9
Treatment effect	-7.49** (2.94)	-8.44*** (2.66)	-8.53*** (2.63)	-9.01*** (2.61)	-8.54*** (2.66)	-6.82** (2.75)	-5.12* (2.85)	-3.92 (3.07)	-4.22 (3.58)

Notes: \*\*\* denotes significance at 1%, \*\* significance at 5%, \* significance at 10%. The coefficients report the quantile differences-in-difference treatment effects for nine quantiles of the test score distribution. The regressions include year dummies and school fixed effects. Bootstrapped standard errors (200 repetitions) adjusted for clustering on the school level are reported in parenthesis.

**Table 5: Estimation Results for Restricted Age Ranges**

Dependent variable: PROEB test scores			
Number of clusters: 1,993			
	(1)	(2)	(3)
Panel A – students in target age range for 4 <sup>th</sup> grade			
Treatment effect	-8.07***	-7.50***	-7.22***
	-2.26	-2.23	-2.25
R-squared	0.01	0.01	0.13
Observations	149,223		
Panel B – repeaters (outside target age range)			
Treatment effect	-3.42	-2.89	-3.29
	(2.50)	(2.50)	(2.42)
R-squared	0.01	0.00	0.08
Observations	88,657		
School fixed effects	yes	yes	yes
Year dummy	yes	yes	yes
School level controls	yes	yes	yes
Peer characteristics controls	no	yes	yes
Individual characteristics controls	no	no	yes

Notes: \*\*\* denotes significance at 1%. The above samples exclude students that are below the target age range. Robust standard errors, adjusted for clustering within schools, are reported in parenthesis. Specification (1) only includes year dummies and school fixed effects, specification (2) additionally controls for a rich set of school characteristics (physical characteristics of the school and the class rooms, teaching material, teacher characteristics, participation in educational programmes etc.), specification (3) additionally controls for peer socio-economic characteristics at the school level and specification (4) also controls for individual characteristics.

**Table 6: Effect of the Adoption of Automatic Promotion on the Socio-economic Composition**

Dependent variables:	Coefficient	Robust standard error
Proportion of white students	0.010	(0.008)
Proportion of mixed students	-0.009	(0.012)
Proportion of black students	0.003	(0.005)
Proportion of Asian students	-0.004	(0.003)
Proportion of indigenous students	-0.000	(0.003)
Mean age (in years)	-0.106***	(0.022)
Mean male students	-0.005	(0.007)
HH wealth index	0.007	(0.025)
Bathroom mean	0.018	(0.009)
TV mean	0.004	(0.007)
Video mean	0.007	(0.012)
Radio mean	-0.010	(0.012)
Fridge mean	0.016**	(0.007)
Freezer mean	-0.014	(0.019)
Washing machine mean	0.012	(0.010)
Car mean	0.011	(0.009)
Computer mean	-0.004	(0.008)
Books mean	0.026	(0.014)

n=1993

Notes: denotes \*\* significance at 5% level, \*\*\* significance at 1% level. All estimates refer to school means or proportions at the school level. All data is taken from the socio-economic questionnaire of PROEB. For each dependent variable the effect is estimated separately in a regression corresponding to equation (1) as  $y_{st} = \beta_0 + d_s + d_t + \gamma d_{st} + \varepsilon_{st}$ . The regression estimates are weighted by school cohort size and include a year dummies ( $d_t$ ) and school fixed effects ( $d_s$ ). Robust standard errors, adjusted for clustering within SREs, are reported in parenthesis. All estimates are weighted by school cohort size.

**Table 7: Effect of the Adoption of Automatic Promotion on Participation in PROEB**

Dependent variable: difference between official student numbers and PROEB participation numbers			
Coefficient	R-squared within	R-squared between	R-squared overall
0.119	0.594	0.020	0.038
(0.976)			

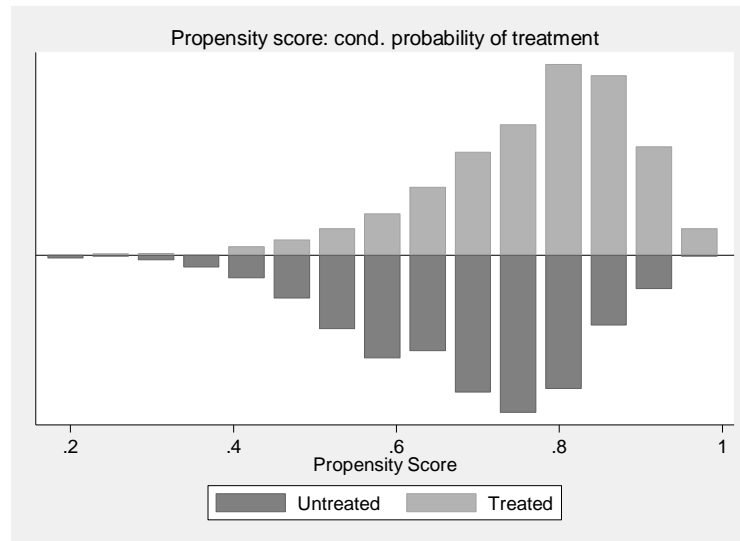
Notes: The coefficient reports the effect of the introduction of automatic promotion on the difference of the number of students according to the school census and the PROEB test. The effect is estimated by a regression corresponding to equation (1) as  $y_{st} = \beta_0 + d_s + d_t + \gamma d_{st} + \varepsilon_{st}$ . The estimates are weighted by school cohort size and include year dummies ( $d_t$ ) and school fixed effects ( $d_s$ ). Robust standard errors adjusted for 46 clusters (on SRE level) are reported in parenthesis.

**Table 8: Effect of Policy Adoption on Student Net Flow**

	(1)	(2)
	Student net inflow up to 1 <sup>st</sup> – 3 <sup>rd</sup> grade	Student net inflow including 4 <sup>th</sup> grade
Coefficient	-0.010 (0.015)	0.079*** (0.014)
R-squared	0.55	0.59

Notes: \*\*\* denotes significance at 1%. The coefficients report the effect of introducing automatic promotion on net flow (including in/outflow due to repetition using data from the school census 2000-2006. A separate regression has been fitted estimating the effect corresponding to equation (1) as  $y_{st} = \beta_0 + d_s + d_t + \gamma d_{st} + \varepsilon_{st}$  for the two models. Model (1) refers to net flows including 2<sup>nd</sup> and 3<sup>rd</sup> grade, model (2) refers to net flows including 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> grade. The regression estimates are weighted by school cohort size and include a year dummy ( $d_t$ ) and school fixed effects ( $d_s$ ). Robust standard errors adjusted for 46 clusters (on SRE level) are reported in parenthesis.

**Figure 1: Histogram for the Propensity Score (year 2003)**



Notes: Figure 1 reports the histogram for the propensity score using all available data on school, teacher and student characteristics from the 2003 school census separately for schools from the treatment and control group.

**Figure 2: Treatment Scheme**

					PROEB test
exam cohort		1 <sup>st</sup> grade	2 <sup>nd</sup> grade	3 <sup>rd</sup> grade	4 <sup>th</sup> grade
Treatment schools	2003	1	1	1	1
	year	2000	2001	2002	2003
	2006	1	0	1	0
	year	2003	2004	2005	2006
Control schools	2003	0	0	1	0
	year	2000	2001	2002	2003
	2006	0	0	1	0
	year	2003	2004	2005	2006

Notes: Testing takes place for all pupils at the end of 4<sup>th</sup> grade. The cohorts are denoted according to the year in which they are tested through PROEB. 1 denotes grades with grade retention, 0 denotes grades with automatic promotion.

**Table 9: Mean Characteristics of Treatment and Control Groups in 2003 and 2006**  
**Panel A: Physical Characteristics and School Programme Participation**

	2003						2006					
	Control		Treatment		P-value	Norm-diff	Control		Treatment		P-value	Norm-diff
n=1993	Mean	SD	Mean	SD			Mean	SD	Mean	SD		
Rural school	0.885	(0.319)	0.890	(0.313)	0.788	-0.011	0.885	(0.319)	0.895	(0.307)	0.544	-0.023
State property	0.918	(0.275)	0.880	(0.325)	0.020	0.089	0.907	(0.290)	0.881	(0.324)	0.106	0.060
Municipal property	0.062	(0.242)	0.093	(0.290)	0.035	-0.082	0.066	(0.249)	0.090	(0.287)	0.098	-0.063
Private property	0.020	(0.141)	0.067	(0.161)	0.414	-0.220	0.024	(0.154)	0.027	(0.163)	0.696	-0.013
School ownership	0.901	(0.298)	0.866	(0.340)	0.040	0.077	0.907	(0.290)	0.881	(0.324)	0.106	0.060
Rented school	0.014	(0.118)	0.020	(0.140)	0.393	-0.033	0.010	(0.100)	0.09	(0.138)	0.164	-0.469
Donated school	0.085	(0.278)	0.114	(0.317)	0.068	-0.069	0.082	(0.275)	0.100	(0.300)	0.261	-0.044
Shared school	0.197	(0.398)	0.238	(0.426)	0.060	-0.070	0.197	(0.398)	0.197	(0.398)	10.00	0.000
Principal office	0.875	(0.331)	0.858	(0.350)	0.324	0.035	0.865	(0.342)	0.874	(0.332)	0.625	-0.019
Admin. office	0.909	(0.287)	0.937	(0.243)	0.036	-0.074	0.950	(0.219)	0.952	(0.214)	0.845	-0.007
Teacher room	0.813	(0.390)	0.811	(0.392)	0.912	0.004	0.827	(0.379)	0.843	(0.363)	0.382	-0.030
School kitchen	0.829	(0.377)	0.824	(0.381)	0.808	0.009	0.821	(0.384)	0.821	(0.384)	0.997	0.000
Refectory	0.374	(0.484)	0.440	(0.497)	0.010	-0.095	0.412	(0.493)	0.464	(0.499)	0.046	-0.074
Food storage	0.839	(0.368)	0.856	(0.352)	0.368	-0.033	0.732	(0.443)	0.718	(0.450)	0.533	0.022
Computer lab	0.237	(0.426)	0.149	(0.356)	0.000	0.159	0.302	(0.460)	0.193	(0.394)	0.000	0.180
Science lab	0.165	(0.372)	0.142	(0.349)	0.205	0.045	0.143	(0.350)	0.120	(0.325)	0.176	0.048
Other lab	0.022	(0.147)	0.021	(0.145)	0.920	0.005	0.020	(0.141)	0.018	(0.133)	0.767	0.010
Toilets outside	0.054	(0.227)	0.094	(0.291)	0.006	-0.108	0.0443	(0.206)	0.070	(0.254)	0.045	-0.079
Toilets inside	0.980	(0.141)	0.967	(0.178)	0.150	0.057	0.992	(0.089)	0.981	(0.136)	0.101	0.068
Freezer	0.899	(0.301)	0.920	(0.271)	0.144	-0.052	0.907	(0.290)	0.914	(0.280)	0.633	-0.017
Filtered water	0.845	(0.362)	0.825	(0.380)	0.299	0.038	0.881	(0.324)	0.858	(0.350)	0.183	0.048

Table 9 continued



Video cassettes	1.757	(1.000)	1.763	(1.190)	0.911	0.000	1.501	(0.936)	1.551	(1.368)	0.444	-0.030
TV sets	2.038	(1.066)	2.025	(1.324)	0.541	0.024	2.107	(1.021)	2.108	(1.353)	0.980	-0.001
Projectors	0.851	(0.544)	0.858	(0.535)	0.815	-0.009	0.867	(0.545)	0.876	(0.530)	0.741	-0.012
Printers	1.790	(2.200)	1.517	(2.040)	0.002	0.110	2.370	(2.295)	1.948	(1.961)	0.000	0.140
Pentium computers	2.757	(4.866)	1.919	(3.915)	0.000	0.151	3.050	(5.053)	2.239	(4.435)	0.001	0.121
386/486 computers	0.656	(2.486)	0.428	(1.469)	0.014	0.079	1.032	(3.109)	0.737	(2.249)	0.022	0.077
Perm. class rooms	10.069	(4.857)	10.060	(4.343)	0.320	0.035	10.651	(5.439)	10.533	(4.403)	0.625	0.017
Prov. class rooms	0.193	(0.581)	0.150	(0.456)	0.092	0.058	0.165	(0.459)	0.172	(0.516)	0.794	-0.010
Class rooms	9.487	(5.188)	9.205	(4.195)	0.223	0.042	10.010	(5.255)	9.760	(4.198)	0.282	0.037
Total staff	49.535	(30.385)	46.612	(25.796)	0.037	0.073	51.553	(31.624)	48.281	(25.749)	0.021	0.080
Teachers	32.177	(20.045)	30.243	(16.885)	0.035	0.074	32.441	(20.651)	30.207	(16.330)	0.014	0.085
Min inc. program	0.598	(0.491)	0.545	(0.498)	0.043	0.076	0.970	(0.171)	0.967	(0.176)	0.778	0.012
TV escola	0.732	(0.443)	0.751	(0.432)	0.401	-0.031	0.495	(0.500)	0.564	(0.496)	0.008	-0.098
Other education TV	0.237	(0.426)	0.304	(0.460)	0.004	-0.107	0.117	(0.321)	0.150	(0.358)	0.062	-0.069
PROINFO	0.199	(0.400)	0.126	(0.332)	0.000	0.140	0.171	(0.377)	0.124	(0.330)	0.009	0.094
State programmes	0.314	(0.465)	0.206	(0.404)	0.000	0.175	0.247	(0.432)	0.225	(0.418)	0.309	0.037
Munic. programmes	0.091	(0.287)	0.102	(0.303)	0.449	-0.026	0.175	(0.380)	0.170	(0.376)	0.787	0.009
School transport	0.527	(0.500)	0.487	(0.500)	0.118	0.057	0.722	(0.448)	0.668	(0.471)	0.025	0.083
Initial enrolment	72.630	(52.326)	69.245	(50.132)	0.197	0.047	38.276	(30.966)	41.202	(32.187)	0.077	-0.066
Classes in 1 <sup>st</sup> grade	2.537	(1.640)	2.359	(1.542)	0.030	0.079	1.682	(1.259)	1.766	(1.308)	0.215	-0.046
Classes in 2 <sup>nd</sup> grade	2.410	(1.420)	2.425	(1.581)	0.854	-0.007	1.793	(1.295)	1.963	(1.324)	0.013	-0.092
Classes in 3 <sup>rd</sup> grade	2.348	(1.428)	2.470	(1.567)	0.125	-0.058	2.163	(1.449)	2.123	(1.420)	0.587	0.020
Classes in 4 <sup>th</sup> grade	2.334	(1.460)	2.469	(1.571)	0.080	-0.063	2.082	(1.289)	2.076	(1.314)	0.919	0.003

Notes: The binary variables of school characteristics and programme participation are coded 0 for not present (no participation) and 1 for present (participation). All data is from the Brazilian school census 2003 and 2006. The p-value is reported from a test on the equality of the mean between the treatment and control groups (independent samples). As the sample size is sufficiently large the result for using a classical t-test or taking into account the binary values and the underlying binomial distribution deliver very similar results. As the group size and with it the variances between the groups differ, approximate t using individual sample variances instead of the pooled variance and Welch's approximation of the degrees of freedom have been used.

The normalized difference is computed as  $norm - diff = \frac{\bar{x}_0 - \bar{x}_1}{\sqrt{s_{X,0}^2 + s_{X,1}^2}}$ , where  $S^2$  denotes the sample variance of  $X_i$ .

**Table 9: Mean Characteristics of Treatment and Control Groups in 2003 and 2006**  
**Panel B: Individual and Family Characteristics at School Level**

	2003						2006					
	Control			Treatment			Control			Treatment		
	Mean	SD	Mean	SD	P-value	Norm-diff	Mean	SD	Mean	SD	P-value	Norm-diff
Male	0.492	(0.500)	0.495	(0.500)	0.414	-0.004	0.490	(0.500)	0.498	(0.500)	0.028	-0.011
Age (in months)	135.43	(14.53)	132.01	(12.72)	0.000	0.177	134.53	(13.73)	132.22	(12.46)	0.000	0.125
% white pupils	0.302	(0.459)	0.333	(0.471)	0.000	-0.047	0.316	(0.465)	0.318	(0.466)	0.479	-0.003
% mixed pupils	0.340	(0.474)	0.354	(0.478)	0.000	-0.021	0.420	(0.494)	0.428	(0.495)	0.021	-0.011
% black pupils	0.118	(0.323)	0.121	(0.326)	0.125	-0.007	0.155	(0.362)	0.146	(0.363)	0.000	0.018
% Asian pupils	0.034	(0.181)	0.034	(0.181)	0.935	0.000	0.045	(0.206)	0.043	(0.203)	0.250	0.007
% indig. pupils	0.042	(0.200)	0.043	(0.203)	0.324	-0.004	0.046	(0.210)	0.046	(0.209)	0.001	0.000
Bathroom	1.267	(0.560)	1.270	(1.291)	0.000	-0.002	1.263	(0.583)	1.276	(0.585)	0.001	-0.016
TV	1.323	(0.792)	1.321	(0.773)	0.762	0.002	1.495	(0.788)	1.487	(0.775)	0.171	0.007
Video	0.370	(0.483)	0.374	(0.484)	0.134	-0.006	0.606	(0.689)	0.605	(0.687)	0.816	0.001
Radio	1.468	(0.796)	1.440	(0.790)	0.000	0.025	1.360	(0.752)	1.341	(0.738)	0.000	0.019
Fridge	0.954	(0.468)	0.978	(0.468)	0.000	-0.036	0.995	(0.493)	1.002	(0.466)	0.054	-0.010
Freezer	1.945	(0.228)	1.940	(0.237)	0.003	0.015	1.937	(0.242)	1.932	(0.252)	0.002	0.014
Clothes washer	0.746	(0.435)	0.781	(0.414)	0.000	-0.058	0.924	(0.601)	0.944	(0.583)	0.000	-0.024
Car	0.621	(0.855)	0.660	(0.862)	0.000	-0.032	0.641	(0.814)	0.672	(0.810)	0.000	-0.027
Computer	0.170	(0.376)	0.176	(0.381)	0.022	-0.011	0.254	(0.435)	0.258	(0.437)	0.229	-0.006
Books	21.867	(27.328)	22.070	(27.612)	0.311	-0.005	20.870	(27.850)	20.471	(27.664)	0.037	0.010
Education (father)	6.318	(19.122)	6.577	(19.822)	0.000	-0.009						
Education ( mother)	6.487	(17.682)	6.651	(18.365)	0.000	-0.006						
Literate (father)	0.877	(0.108)	0.887	(0.100)	0.000	-0.068						
Literate (mother)	0.905	(0.086)	0.910	(0.082)	0.029	-0.038						

Notes: All data is taken from the socio-economic questionnaire of PROEB 2003 and 2006. Information on educational background and literacy of parents is only available in the 2003 questionnaire. The p-value is reported from a test on the equality of the mean between the treatment and control groups (independent samples). As the sample size is sufficiently large the result for using a classical t-test or taking into account the binary values and the underlying binomial distribution deliver very similar results. As the group size and with it the variances between the groups differ, approximate t using individual sample variances instead of the pooled variance and Welch's approximation of the degrees of freedom have been used.

The normalized difference is computed as  $norm - diff = \frac{\bar{x}_0 - \bar{x}_1}{\sqrt{s_{X,0}^2 + s_{X,1}^2}}$ , where  $S^2$  denotes the sample variance of  $X_i$ .

**Table 10: Mean Pre-intervention School Characteristics in 1997**

1997						
n=1,993						
	Control		Treatment		P-value	Norm-diff
	Mean	SD	Mean	SD		
Repetition rate 1 <sup>st</sup> grade	0.08	(0.108)	0.08	(0.106)	0.98	-0.001
Repetition rate 2 <sup>nd</sup> grade	0.16	(0.208)	0.17	(0.200)	0.37	-0.034
Repetition rate 3 <sup>rd</sup> grade	0.05	(0.053)	0.05	(0.143)	0.63	-0.014
Repetition rate 4 <sup>th</sup> grade	0.06	(0.072)	0.06	(0.093)	0.21	-0.044
Class size (grades 1-4)	32.249	(5.871)	31.127	(6.480)	0.000	0.128
Student-teacher ratio (grades 1-4)	20.480	(3.878)	20.129	(4.124)	0.244	0.062

Notes: All data is from the Brazilian school census 1997. The p-value is reported from a test on the equality of the mean between the treatment and control groups (independent samples). As the sample size is sufficiently large the result for using a classical t-test or alternatively taking into account the binary values and the underlying binomial distribution deliver very similar results. As the group size and with it the variances between the groups differ, approximate t using individual sample variances instead of the pooled variance and Welch's approximation of the degrees of freedom have been used.

The normalized difference is computed as  $orm - diff = \frac{\bar{x}_0 - \bar{x}_1}{\sqrt{s_{X,0}^2 + s_{X,1}^2}}$ , where  $S^2$  denotes the sample variance of  $X_i$ .

**Table 11: Responses of Students, Parents and Teachers**

Dependent variable:				
	Students doing homework (1)	Parents help with homework (2)	Teacher assigns homework (3)	Teacher corrects homework (4)
	-0.014* (0.008)	0.022*** (0.008)	-0.002 (0.003)	-0.002 (0.009)
Observations	217,253	212,647	220,087	215,809
R-squared	0.003	0.041	0.001	0.003

Notes: \*\*\* denotes significance at 1%, \* denotes significance at 10%. Robust standard errors, adjusted for clustering within schools are reported in parenthesis. The coefficients report the effect of the introduction of automatic promotion behavioural responses of students, parents and teachers. The effects are estimated by a regression corresponding to equation (1) as  $y_{ist} = \beta_0 + d_s + d_t + \gamma_{st} + \epsilon_{ist}$ . The binary outcome variables were constructed using consistent information from the socio-economic questionnaire of PROEB 2003 and 2006. The dependent variable in column (1) reports change in fraction of students always doing their homework (mean 0.706), in column (2) the change in fraction of students always receiving help from their parents with their homework (mean 0.652), in column (3) the change in fraction of teachers assigning homework (mean 0.981), and in column (4) the change in fraction of teachers always correcting homework of their students (mean 0.767).

**Table 12: Sensitivity of Estimates to Individual and Peer Age Controls for Different Age Ranges**

Dependent variable: PROEB test scores								
Panel A – all pupils								
	Peer controls		Peer and individual controls				Individual controls	
	excl. peer age	incl. peer age	excl. ind. and peer age	incl. ind. age, excl. peer age	incl. peer age, excl. ind. age	incl. peer and ind. age	excl. ind. age	incl. ind. age
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment effect	-5.65*** (1.98)	-5.67*** (1.98)	-6.02*** (1.98)	-6.47*** (2.04)	-5.78*** (1.99)	-6.24*** (2.04)	-6.76*** (1.96)	-6.87*** (2.08)
R-squared	0.01	0.01	0.12	0.17	0.12	0.17	0.12	0.16
Observations: 244,081								
Panel B – students in target age range for 4 <sup>th</sup> grade								
	Peer controls		Peer and individual controls				Individual controls	
	excl. peer age	incl. peer age	excl. ind. and peer age	incl. ind. age, excl. peer age	incl. peer age, excl. ind. age	incl. peer and ind. age	excl. ind. age	incl. ind. age
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment effect	-7.47*** (2.23)	-7.50*** (2.23)	-7.30*** (2.25)	-7.46*** (2.25)	-7.11*** (2.25)	-7.22*** (2.25)	-7.73*** (2.26)	-7.96*** (2.26)
R-squared	0.01	0.01	0.13	0.13	0.13	0.13	0.12	0.12
Observations: 149,223								

Notes: \*\*\* denotes significance at 1%. All estimates include controls for school characteristics (physical characteristics of the school and the class rooms, teaching material, teacher characteristics, participation in educational programmes etc.) The specifications (1) include additionally controls for peer socio-economic characteristics, specifications (2) control for per and individual characteristics, specifications (3) control for individual characteristics. The row below specifies further controls for individual and peer age in the estimation.

**Table 13: Effect of Policy Adoption on 3rd Grade Repetition Rate for Treatment Schools**

	3 <sup>rd</sup> grade repetition
Coefficient	0.0428*** (0.0099)
R-squared	0.019

Notes: \*\*\* denotes significance at 1%. The coefficient reports the effect of introducing automatic promotion on 3<sup>rd</sup> grade repetition rate for the cohort of interest of treatment schools.

**Table 14: Linear Probability Model of Assignment to Treatment**

	(1)		(2)	
School provides all years of fundamental edu.	0.061**	(0.030)	0.067**	(0.027)
School characteristics				
rural school	0.004	(0.039)	0.001	(0.036)
proper school building	0.070	(0.099)	0.100	(0.099)
church building	-0.152**	(0.070)	-0.128	(0.076)
teacher home	-0.156	(0.175)	-0.289	(0.170)
building of other school	0.013	(0.061)	-0.016	(0.053)
farm building	-0.111	(0.120)	-0.037	(0.113)
other building	-0.107	(0.059)	-0.119**	(0.053)
state property	-0.271	(0.240)	-0.130	(0.233)
municipal property	-0.380	(0.233)	-0.210	(0.229)
private property	-0.348	(0.237)	-0.227	(0.227)
school property	-0.061	(0.070)	-0.076	(0.059)
rented property	0.034	(0.109)	0.045	(0.097)
shared school	-0.038	(0.023)	-0.052**	(0.021)
principal office	0.007	(0.033)	-0.023	(0.030)
secretarial office	-0.067	(0.042)	-0.083**	(0.040)
school library	-0.008	(0.030)	0.016	(0.027)
reading room	-0.029	(0.052)	-0.087	(0.046)
teacher room	-0.004	(0.030)	-0.001	(0.027)
video library	-0.031	(0.041)	-0.050	(0.039)
TV room	0.006	(0.024)	0.021	(0.022)
school kitchen	0.016	(0.029)	-0.018	(0.027)
cafeteria	0.008	(0.024)	0.011	(0.021)
refectory	-0.045**	(0.021)	-0.023	(0.020)
stationary	-0.025	(0.024)	0.003	(0.022)
computer laboratory	0.017	(0.042)	0.012	(0.039)
science laboratory	0.011	(0.031)	0.003	(0.029)
other laboratory	0.018	(0.061)	-0.005	(0.053)
toilets outside school	-0.088**	(0.036)	-0.053	(0.033)
toilets inside school	0.023	(0.055)	-0.013	(0.052)
toilets ready for special needs	0.108	(0.083)	0.144	(0.076)
school ready for special needs	-0.062	(0.071)	-0.022	(0.062)
industrial oven	0.021	(0.089)	0.056	(0.088)
home oven	-0.022	(0.030)	-0.009	(0.029)
wood oven	0.052	(0.048)	0.016	(0.037)
freezer	-0.052	(0.038)	-0.056	(0.034)
filtered water	0.051**	(0.026)	-0.007	(0.024)
internet access	0.017	(0.033)	-0.012	(0.030)
public energy supply	-0.044	(0.247)	-0.049	(0.162)
solar energy	-0.229***	(0.071)	-0.100	(0.086)
using 220 volt	-0.020	(0.040)	-0.008	(0.037)
using 110 & 220 volt	-0.063**	(0.029)	-0.010	(0.029)
public water supply	-0.025	(0.089)	0.061	(0.077)
artesian well	-0.082	(0.086)	0.003	(0.076)
cistern water	-0.003	(0.073)	-0.021	(0.063)
no running water	-0.093	(0.250)	-0.075	(0.267)

Table 14 continued	spring water	-0.060	(0.105)	-0.034	(0.093)
	public sewerage	0.055	(0.303)	0.125	(0.240)
Number of	septic tank	-0.070	(0.303)	0.098	(0.241)
	no sewerage	0.121	(0.319)	0.135	(0.250)
	video tapes	-0.012	(0.014)	-0.010	(0.013)
	TV sets	-0.006	(0.012)	-0.004	(0.012)
	printers	-0.001	(0.009)	0.002	(0.008)
	overhead projectors	-0.043**	(0.021)	-0.015	(0.020)
	Pentium computers	0.008	(0.004)	0.004	(0.004)
School size	386/486 computers	0.013**	(0.006)	0.013**	(0.006)
	permanent class rooms	-0.015	(0.009)	-0.007	(0.008)
	provisory class rooms	0.019	(0.023)	0.015	(0.022)
	class rooms in school	0.017	(0.010)	0.013	(0.009)
	class rooms away from school	0.030	(0.015)	0.023	(0.013)
	student enrolment 1st year	-0.000	(0.000)	0.000	(0.000)
	total number of staff	0.000	(0.003)	0.002	(0.003)
Programme participation	total number of teachers	0.001	(0.004)	-0.001	(0.004)
	number of teachers in classes 1-4	-0.003	(0.002)	-0.003	(0.002)
	Minimum Income Programme	0.012	(0.020)	0.004	(0.022)
	TV escola	-0.007	(0.024)	0.012	(0.023)
	Education TV	-0.040*	(0.022)	-0.013	(0.021)
	Parameters in Action	-0.045	(0.028)	-0.025	(0.027)
	FNDE	0.005	(0.025)	-0.025	(0.025)
	Ouvebem	-0.061**	(0.028)	-0.025	(0.026)
	Reabvis	-0.025	(0.027)	-0.007	(0.025)
	School Transport Programme	0.074***	(0.027)	0.021	(0.027)
	National Library Programme	-0.025	(0.021)	-0.006	(0.019)
	State computer programme	0.038	(0.062)	0.043	(0.059)
	Municipal computer programme	0.213	(0.179)	0.153	(0.170)
	Proinfo	0.023	(0.039)	0.058	(0.036)
	other state programme	0.104***	(0.025)	0.008	(0.024)
	other municipal programme	-0.039	(0.032)	-0.068**	(0.031)
	Free School Lunch	0.047	(0.170)	0.007	(0.181)
	Free Public School Transport	-0.035	(0.024)	0.007	(0.022)
	Constant	0.634	(0.504)	0.807	(0.440)
Observations		1993			
SRE dummies		No		Yes	
R-squared		0.09		0.28	

Notes: \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%, standard errors are in parenthesis. Specification (2) includes regional school administration dummies (SRE). Most of the physical characteristics describing the schools are indicator variables on the presence at school. Similarly, indicator variables inform about participation in education programmes. The programme *Parameters in Action* is a federal programme for the professional development of teachers; *FNDE* denotes a maintenance and development programme for education by the National Fund for the Development of Education, *Ouvebem* is a national campaign for the importance of the sense of hearing, *Reabvis* is a national campaign on visual rehabilitation, *PROINFO* is a federal computer literacy programme. This table reports the coefficients from two specifications of a linear model of the effect of school characteristics on the probability for treatment. The dependent variable is an indicator that equals zero for being in treatment group and equals 1 for being in the control group. None of the coefficients of the linear model produces values outside the unit-interval and a logit specification delivers very similar results to the linear specification diminishing doubts on the suitability of the linear specification (not reported).



## 2. Class Assignment and Peer Group Effects

### 2.1 Introduction

The question of whether the composition of the peer group matters for the outcome of an individual member of the group has received considerable attention in numerous contexts where social interactions can be present. Peer effects have been studied in the context of schools, universities, workplaces, neighbourhoods and prisons, among others.<sup>32</sup> Due to the natural grouping of students into schools and classrooms, and the potential for education policies to affect the peer group composition, peer effects in education have received extensive attention by economists. Recent work goes beyond linear-in-means specifications and points to the potential relevance of the distribution of peer characteristics in explaining group effects (Hoxby and Weingarth 2006, Lyle 2009).

The identification of group effects is challenging due to conceptual problems as well as data limitations. In the education sphere, for example, an identification strategy for peer effects needs to address potential endogenous selection of students into schools and classes. With selection into groups, unobserved characteristics such as ability, parental support or students' effort are likely to be correlated among peers, and educational outcomes are therefore correlated within the peer group even in the absence of externalities. In addition, the analysis needs to deal with separating peer effects from

---

<sup>32</sup> Recent studies include Mas and Moretti (2009) on productivity effects for supermarket cashiers; Bandiera, Barankay and Rasul (2010) on social networks and worker productivity in farm production; Bayer, Hjalmarsson and Pozen (2009) on the effect of juvenile offenders' serving time on other's subsequent criminal behaviour to name just a few. Studies on peer effects in education include Hoxby (2000) for gender and race peer effects; Hanushek et al. (2003) provide a framework for estimating peer effects trying to overcome omitted variables and simultaneous equation biases; Duflo, Dupas and Kremer (2010) provide evidence from a randomized experiment in Kenya; Lavy, Paserman and Schlosser (2008) on ability peer effects and potential channels; Lavy, Silva and Weinhardt (2009) on distributional effects of ability peer effects; Lavy and Schlosser (2011) on gender peer effects and their operational channels; Zimmerman (2003) and Sacerdote (2003) for peer effects in college education; Angrist and Lang (2004) for peer effects on racial integration and Ammermueller and Pischke (2009) for a cross-country comparison of peer effects at primary school level. Student tracking, school choice, busing, admission policies, class formation, repetition policies, and residential location decisions are relevant policy issues that can change the peer composition in schools and classrooms (Zimmerman 2003 and Hanushek et. al 2003).

common shocks to the peer group, such as differential educational and teacher inputs, and it needs to account for simultaneous determination of student and peer achievement (Manski 1993, Hanushek et al. 2003).

Randomized experiments are the first choice for overcoming the selection problem and there are a number of recent applications in this area (see Duflo, Dupas and Kremer (2011) on ability grouping in primary schools, and Whitmore (2005) on gender peer effects in higher education, Carrell et al. 2009 on peer effects in higher education and Carrell et al. 2013 on the endogenous formation of peer subgroups). Empirical strategies that exploit natural experiments, such as conditional random assignment of college roommates by Zimmerman (2003) and Sacerdote (2003), or the idiosyncratic variation in the composition of a given cohort over time have also been used (Hoxby, 2000, Ammermueller and Pischke 2009, Gibbons and Telhaj 2012, Ohinata and van Ours 2013). There is little experimental or quasi-experimental evidence that overcomes the identification problems of peer group effects in primary or secondary education and even less evidence that specifically considers distributional features of peer groups that might affect educational achievement.

Group heterogeneity has not received much attention in the literature on peer effects. It has though been addressed in the literature on tracking (also referred to as streaming), where students are separated by academic ability into schools or classes.<sup>33</sup> Some recent research on the effects of tracking that addresses the endogeneity of tracking decisions, finds that tracking may benefit equally students from lower and higher achievement tracks. Figlio and Page (2002) show that tracking may actually help low-ability students without proposing a specific mechanism for this effect and Zimmer (2003) presents quasi-experimental evidence that a negative direct peer effect for low-

---

<sup>33</sup> There is an extensive pedagogic literature on age, ability grouping, and academic tracking. See Robinson (2008), Adams-Byers, Squiller Whitsell and Moon (2004), and Betts and Shkolnik (1999) for some recent examples. Kremer (1997) provides an economic model of sorting.

achieving students is offset by the positive effects of achievement targeted instruction. Duflo, Dupas and Kremer (2011) use quasi-experimental assignment of pupils to classes to study the effect of tracking students on initial achievement among Kenyan primary school students. They find persistent positive effects across the achievement distribution of tracking students in a higher and a lower ability class. They attribute this effect mainly to teacher effort and the choice of target teaching level given the particular incentives for teachers in Kenyan schools, and the better match of the instruction level due to reduced heterogeneity in ability in the classrooms. Their results are matched by the findings of Zimmer (2003) and Hoxby and Weingarth (2006) who show that students in more homogenous classes benefit from more tailored instruction. De Giorgi, Pellizzari and Woolston (2010) provide evidence on the effect of class heterogeneity on academic achievement and labour market outcomes in the setting of higher education. They find that the effect of the peer distribution on student performance is non-linear and appears to be inverse U-shaped with respect to the dispersion of gender and ability in the group.

This chapter provides a novel method to identify peer effects in education and quasi-experimental evidence from exogenous variation in peer group membership by using an assignment mechanism of students into classes which provides the basis for a regression discontinuity (RD) design. Brazilian primary school students are typically allocated to classes based on their relative age in the cohort. Using the age rank as a continuous assignment variable, this rule creates a discontinuity in the allocation to a class (peer group) for students close to the class size cap of the relatively younger class. I exploit this rule to compare outcomes of students at the margin of being assigned to an older versus a younger group in schools with two classes per cohort. The present setting differs from the settings in most of the literature where identification relies on small

differences in peer characteristics from idiosyncratic variation across cohorts or from variation based on random assignment.<sup>34</sup>

Using two-stage-least squares to estimate the discontinuity in a fuzzy RD setting, I find strong evidence for sizeable peer group effects. I estimate a negative effect from being in the relatively older class on maths test scores among students in fifth grade of around half of a standard deviation.

A major challenge in the present context is that the discontinuity cut-off - i.e. the size of the younger class - is potentially endogenous. If students are strategically allocated to classes based on their latent outcomes, variation in outcomes around the threshold is not ‘as good as random’ and differences in outcomes between those on the right and on the left of the cut-off do not provide a consistent estimate of the parameter of interest (Lee and Lemieux 2010). In the chapter, though, I argue that assignment to the groups is largely predetermined (in 1<sup>st</sup> grade) and I find no evidence, based on a large array of observable covariates, of non-random sorting around the cut-off.<sup>35</sup>

Because I have data on more than 350 schools, I am able to estimate a separate parameter for each school and relate the magnitude of the estimated coefficient to differences in exogenous class characteristics across schools. This strategy allows me to learn about which observable differences across classes, if any, drive the estimated gap in the attainment between barely eligible and barely ineligible pupils. Because, in Brazil, as in many other low- and middle-income countries, grade repetition is widespread, older classes tend typically to display larger variation in age. I find evidence that the estimated group effect may be due to the difference in the age dispersion across classes.

---

<sup>34</sup> Carrell et al. 2013 is an exception to this.

<sup>35</sup> Appendix A2 provides information on the initial assignment of students and the transition from one grade to the following grade.

The remainder of this chapter is organized as follows: Section 2.2 briefly describes the Brazilian educational system and the educational system in the state of Minas Gerais, which is the focus of this study. Section 2.3 describes the data. Section 2.4 presents the assignment mechanism of students to classes and introduces the identification strategy. Section 2.5 presents the main results and Section 6 presents tests for non-random sorting and for correlated effects. Section 2.7 gives an interpretation of the peer group estimates and Section 2.8 concludes.

## **2.2. The allocation of students into classes in Minas Gerais**

In the public schools of Minas Gerais, which are the focus of this analysis, “normal” class size is set at 25 students per class.<sup>36</sup> When enrolment per grade is above 25 pupils, the school administration needs to make a choice on how to assign students to classes before the start of the school year. As, unlike innate ability or behavioural characteristics, age of students at the point of enrolment in first grade can be easily observed by school administrators, age sorting provides a convenient and widely used way of grouping students utilising observable characteristics at the time of entry into primary school.<sup>37</sup>

Students who progress regularly typically remain in their original class throughout primary school, so that, other than because of migration between schools and drop-out, assignment to classes is largely predetermined in first grade and not based on students’ observable characteristics other than age.<sup>38</sup> Obviously, grade repetition may potentially lead to changes in the original class assignment. Although grade repetition has been

---

<sup>36</sup> Law 16.056 of 24<sup>th</sup> April 2006 limits class size to 25 students in the initial years of primary education (1<sup>st</sup>-5<sup>th</sup> grade) in all public schools in Minas Gerais. Exceptions are theoretically only allowed under special circumstances and during the transitional period of the introduction of the law. [http://crv.educacao.mg.gov.br/sistema\\_crv/banco\\_objetos\\_crv/%7B103FA0DB-B47A-4E66-A719-402B21F94D5B%7D\\_lei%2016056%202006.pdf](http://crv.educacao.mg.gov.br/sistema_crv/banco_objetos_crv/%7B103FA0DB-B47A-4E66-A719-402B21F94D5B%7D_lei%2016056%202006.pdf)

<sup>37</sup> Grouping students according to their age may in fact at least partially coincide with grouping according to ability, as ability is likely to be correlated with age at time of primary school enrolment. See Cascio and Whitmore Schanzenbach (2007) and Angrist and Krueger (1991) for a discussion of student age and educational outcomes.

<sup>38</sup> The appendix A2 provides more information on the initial assignment of students.

reduced by the introduction of automatic grade promotion in Minas Gerais which chapter one is concerned about, Table 15 reveals that there still exists a substantial number of students who have repeated at least one school grade.<sup>39</sup> Grade repeaters in first grade are, consistent with an assignment rule based on the age ranking of students in the cohort, usually allocated to the older class when repeating the grade in the following year. In succeeding grades, repeaters regularly are allocated to the older class as well. The propensity for repetition in subsequent grades is nevertheless also higher in the older classes, so that the in- and outflow of students into the classes largely cancel out each other and class size is hence largely unaffected by repetition.

### 2.3 Data and descriptive statistics

For the purpose of this analysis, I use standardized test scores in mathematics of primary school students in public schools in Minas Gerais. Educational standards in Minas Gerais are among the highest for the Brazilian states.<sup>40</sup> The primary source of data in this study is of PROEB (the same source of data used in the first chapter), which provides maths test scores at the pupil level for all students in 5<sup>th</sup> grade in the state.<sup>41</sup> I use the data for 2007, as this is the only year that contains detailed information on students' age. The test is carried out at all public schools in the state and test scores are standardized to a mean of 500 with a standard deviation of 100. Participation is compulsory at school and individual levels, confirmed by a high student participation rate (93%). Surveyed pupils also answer a detailed socioeconomic questionnaire, which includes information on sex, month and year of birth, racial background and information on the socioeconomic background of the family.

---

<sup>39</sup> This is due to the fact that grade retention is still possible in third grade and due to excessive school absenteeism.

<sup>40</sup> In the nation-wide school evaluation system of SAEB, 2005 mean maths performance of pupils from Minas Gerais is clearly above the Brazilian average, ranking first among Brazilian states ([http://www.inep.gov.br/salas/download/prova\\_brasil/Resultados/Saeb\\_resultados95\\_05\\_UF.pdf](http://www.inep.gov.br/salas/download/prova_brasil/Resultados/Saeb_resultados95_05_UF.pdf)).

<sup>41</sup> In this chapter I use the 2007 wave which tested 5<sup>th</sup> graders, due to the change of the lengths of primary school from eight years to nine years.

In the following, I restrict the sample to schools with only two classes. This ensures that enough variation is available to identify sizeable group effects for students around the cut-off point, in particular with respect to variation in distributional features of the class composition.<sup>42</sup>

The data comprises 16,031 students from 363 public primary schools. Table 15 presents summary statistics for these data. The average age of students on the test day is 11.27 years, which is about nine months above the normal age for this grade. This age-grade mismatch is due to a combination of late enrolment and grade repetition. Students at these schools are overwhelmingly from deprived socioeconomic family backgrounds and 47% of the families of the students at these schools are recipients of *Bolsa Família*, the Brazilian conditional cash transfer programme for poor and very poor families, compared with around 25% in the total population.<sup>43</sup>

PROEB also includes headmaster and teacher questionnaires. The headmaster questionnaire includes questions on individual characteristics of the headmaster, such as age, sex and educational background and questions on school characteristics and pedagogic strategy at the school. The teacher questionnaire includes questions on individual characteristics, as well as statements on the students in class.

For part of the analysis, I complement the analysis with data from the 2007 School Census, which comprises detailed information on school characteristics for all primary schools in Brazil. The data appendix provides detailed information on the data

---

<sup>42</sup> The focus on schools with two classes ensures that school administrators cannot establish *special classes* that do not follow the general assignment mechanism. With more than two classes the school administration may resort to forming separate classes in which students with specific characteristics are grouped, such as grade repeaters, and are separated from the other students in the cohort. As these *special classes* tend to be rather small, measures of age variation are also more susceptible to outliers (Lyle 2009).

<sup>43</sup> Families are eligible for *Bolsa Família* if per capita family income is not above R\$ 120 per month ('moderately poor') (US\$ 63 at 1<sup>st</sup> June 2007) and receive monthly R\$ 20 per child under the condition of regular school attendance and participation in vaccination campaigns. Families below a per capita income R\$ 60 ('extremely poor') receive an additional basic family allowance of R\$ 62. See <http://www.mds.gov.br/bolsafamilia/> and Lindert et al. (2007) for details.

sources and the variables used. Summary statistics from the census for the schools used in this analysis are presented in Table 22.

## 2.4 Empirical strategy

The identification strategy used in this chapter exploits the discontinuity in the assignment rule of students in schools with two classes. The treatment assignment mechanism is based on the value of an observed and continuous variable, the age rank ( $n$ ) of the individual student in each school, in such a way that the probability of receiving treatment is a discontinuous function of that variable at the class size cap  $\bar{N}$ , the size of the younger class.

Consider a simple reduced-form model of school achievement  $f$

$$Y_{is} = \delta_0 + \delta_1 T_i + f(n) + \varepsilon_i \quad (3)$$

where  $Y_{is}$  denotes the outcome variable maths test score for individual  $i$  in school  $s$ , and  $T_i$  is the treatment indicator that takes a value of 0 for individuals in the younger class and 1 for individuals in the older class,  $\varepsilon_i$  is an individual unobserved error component, I ignore at this stage any covariates one might want to include in the specification to reduce sampling variability in the estimator. Educational achievement measured in terms of test scores is assumed to depend on a smooth function  $f(\cdot)$  of the student's age rank, and on being in either the younger or older class indicated by  $T_i$ . I employ two-stage least squares to estimate  $\delta_1$ , the coefficient of interest, using the discontinuity at the class cap as an instrument for treatment  $T_i$  (being in the older class).

In a first stage-equation, I assume that  $T_i$  is a function of age rank of students in the school cohort and a dummy  $D_{is}$  for being above or below the school-specific discontinuity point  $\bar{N}$  given by the maximum class size rule:

$$T_i = \gamma_1 + \gamma_2 D_{is} + f(n) + \nu_i \quad (4)$$



where  $v_i$  is an error component.

For identification of the class effect  $\delta_1$ , a continuity assumption needs to be satisfied, such that student achievement varies continuously with the forcing variable of the age rank in the cohort, outside of its influence through treatment  $T_i$  (Lee and Lemieux 2010), so that assignment to either side of the discontinuity threshold is as good as random. In other terms, identification of the treatment effect relies on the assumption that just below and above the known cut-off point individuals are similar in observable and unobservable characteristics, other than being in different classes. In this way, the proposed RD strategy allows me to circumvent confounding effects induced by non-random sorting of individuals across groups that plagues the literature on spillover effects.

## 2.5 Empirical results

Before presenting the regression analysis, it is useful to show the raw data. The upper graph of Figure 3 plots standardized local averages of the class rank (1 or 2, to denote respectively group 1 or 2) in one month bins, where the age rank has been centred on the cut-off point of zero. Local linear regression fits using a rectangular kernel with a bandwidth of 3 months are superimposed. The discontinuity in the average class rank at the cut-off point is evident and the size of the discontinuity in the probability of treatment conditional on the age rank is around 0.5. The estimated increase in the rank is less than one, as not all schools choose to allocate students into homogenous classes. It appears that smaller schools deviate from this rule, but other than this, I find little systematic association between the probability of using the age-ranking rule and observable school and pupils characteristics.<sup>44</sup>

---

<sup>44</sup> In the Appendix I estimate a linear probability model, where the dependent variable is a binary variable with a value of 1 if student assignment is based on the age ranking and zero otherwise. I use the rich information in the two datasets on school, headmaster, teacher and students' characteristics to learn about

In panel B of Figure 3, I plot local averages of maths test scores and the local linear regression lines on both sides of the cut-off point. The data show a very clear fall in maths test scores: the oldest pupil in the younger class shows an average attainment in maths that is 0.2 of a standard deviation higher than that of the younger pupil in the older class. Hence, Figure 3 suggests that being assigned to the older class significantly harms learning outcomes.

Table 14 presents the first-stage estimates for the size of the discontinuity in mean class rank, the OLS estimates for the size of the discontinuity in test scores at the discontinuity point and the 2SLS estimates for the causal effect of crossing the cut-off point from the younger class to the older class. All specifications include school-fixed effects that account for observed and unobserved differences across schools which are common across classes. Standard errors are heteroskedasticity consistent and adjusted for clustering at the school level. Column (1) presents the estimates for the models including only a quadratic polynomial in age rank. Column (2) includes controls for the whole set of predetermined individual and family characteristics. The estimates of column (3) include teacher characteristics in addition to the other covariates.

The top panel of Table 16 presents estimates for the first stage regressions, where the dependent variable is 1 for students being in the older class and zero otherwise. The

---

the determinants of the allocation rule. Specifically I estimate the following linear model:  $Y = \beta_0 + \beta_1 S + \beta_2 D + \beta_3 T + \beta_4 P + u$ , where  $Y$  takes a value of 1 for an allocation rule that sorts students into homogenous age classes and a value of 0 otherwise.  $S$  denotes school characteristics,  $D$  headmaster characteristics,  $T$  teacher characteristics,  $P$  mean characteristics of pupils in the cohort and  $u$  an idiosyncratic error term. Table A2 reports the coefficients from the estimated model. Only few variables show a statistically significant effect at conventional levels of significance: cohort size, the existence of a headmaster's office, and the headmaster being of an Asian or Indigenous background and the mean number of fridges in student's families. With a larger cohort size, administrators are inclined to choose homogenous age sorting. The socioeconomic composition of students in the cohort and mean teacher characteristics do not seem to play a role in the choice of the assignment rule of students to classes. Other coefficients, such as the existence of a copy machine, the headmaster being of Indigenous background, the proportion of Asian students and the mean number of fridges in the student households are only statistically significant at the 10% level. In sum, there is little evidence of a systematic choice of the allocation rule based on headmaster, teacher or student characteristics.

estimates for the size of the discontinuity range between 0.451 and 0.467, similar to the observed discontinuity in panel A of Figure 3.

The middle panel of Table 16 reports the reduced form estimates from an OLS regression with maths test scores as the dependent variable on a dummy equal to 1 for being to the right of the threshold. Column 1 reports the raw estimate of the discontinuity of maths test scores at the cut-off point.

The bottom panel of Table 16 reports the two-stage-least squares estimates for the class peer effects using the same specifications as for the OLS estimates in panels A and B. The size of the class peer effect, without further controls, is around 0.57 of a standard deviation in maths test scores and significant at the 1% level.

Under the identifying assumptions outlined in the previous section, the results can be interpreted as the causal effect on individuals whose treatment status changes, i.e. who were to switch from the younger class to the older class as the value of  $n$  changes from just below  $\bar{N}$  to just above  $\bar{N}$ .

To acquire some understanding of the distribution of effects across schools, I estimate school-specific discontinuities in maths test scores. As differences of mean peer variables between classes differ across schools, treatment also differs in respect of the composition of the peer class environment. Figure 4 plots the kernel density estimates of the school-specific discontinuities and shows the relatively symmetric distribution of effects around a peak of about -50. I will return to heterogeneous effects across schools in Section 7.

Table 17 presents the RD estimates for wider intervals of the discontinuity sample around the cut-off point and different orders of the polynomial terms included in the regressions as a first robustness check. Rows 1 and 2 are the estimates of the RD without any further controls, rows 3 and 4 are the estimates including the full set of controls including individual, family and teacher characteristics. The estimates do not

reveal any substantial sensitivity with respect to the choice of the order of the polynomial. Replacing the quadratic by a cubic term leaves the estimates virtually unchanged. Increasing the range of observations used for the estimation also does not alter the estimates for the treatment effect in any significant way.

## **2.6 Tests for non-random sorting**

As already outlined, there are obvious threats to the identification assumption. Public knowledge of the allocation mechanism and the alleged penalty associated with treatment may invalidate the continuity assumption required for consistency of the RD estimator if the forcing variable is subject to manipulation by optimizing agents (McCrary 2008). In the present context, there is potential for manipulation of the forcing variable by two sets of agents involved, the parents of the school children and school administrators. If either parents or school administrators are able to manipulate the assignment of a student precisely around the cut-off point, the ‘as good as random’ assignment may fail.<sup>45</sup>

In the parents’ case, a threat to the identification strategy arises from parents exerting pressure on school administrators to assign their child to the younger class at the time of initial enrolment or at a later stage. For the case of students close to the cut-off point, if the ability of parents to exert pressure and move their child to the younger class is systematically related to other unobserved determinants of maths achievement (e.g. the home learning environment or the support the student receives) this may invalidate the assumptions of the RD design.

Similarly, the school administration might manipulate class size so to move the youngest student in the older class to the younger class, or vice versa, based on some

---

<sup>45</sup> McCrary (2008) suggests a test for the failure of the random assignment assumption by inspecting for a discontinuity in the density of the forcing variable around the discontinuity point. As the forcing variable in the present case is uniformly distributed due to its nature as a relative rank, this test will not be informative in this analysis.

characteristics that are not necessarily observable to the econometrician and that are correlated with outcomes. In this case, the cut-off point would simply be shifted by one rank upwards/downwards. In reality this is unlikely to happen, as the allocation of students is decided before classes start at first grade, so that the school administration has no information on ability, race or socioeconomic background of the student other than administrative information such as age or sex that is to be found in the documents necessary for enrolment, such as a birth certificate.

In all cases, if manipulation occurred, whether due to schools or parents' pressure, pre-determined characteristics of students and their families would presumably no longer be balanced on either side of the discontinuity (van der Klaauw 2002).

In the following I use a very rich array of information from the student questionnaire to formally test for the balancing properties of pre-determined student characteristics across the cut-off point. Figure 8 provides a graphical analysis of the balancing properties of baseline covariates by plotting local averages for the covariates and local linear regression fits separately on both sides of the threshold. In Figure 8 (part 1), the graphs in columns 1 and 3 plot the individual level probability of being a girl and the probability of self-identifying with different ethnic groups. The fraction of girls reduces smoothly with the age rank. The fraction white, Asian or indigenous students in the class does not reveal any discontinuity at the threshold, while the fraction of mixed and black students show a minor positive increase at the cut-off point. The average number of months repeated before also does not reveal a discontinuity, but different slopes of the local linear regression fits are apparent, these being induced by the different distribution of repeaters in the two classes. Columns 1 and 3 of Figure 8 (continued) present the same graphs for a wide range of predetermined socioeconomic characteristics. These variables appear well balanced on both sides of the cut-off point and there is no indication of a discontinuity in the means of these characteristics at the

cut-off point. Among two additional proxies for the socioeconomic status of the family, the number of domestic workers employed and the fraction of families receiving *Bolsa Família*, only the latter shows a small difference around the threshold.

In a formal analysis, I estimate all predetermined characteristics of students using the same specification as for the main estimates in Table 16. Table 18 reports the RD estimates for these variables. Only the estimate for the probability of being a black student is significant at the 5% level.<sup>46</sup> None of the other household socioeconomic characteristics reveal a statistically significant difference at the threshold and most coefficients are small, confirming that the balancing properties of these predetermined characteristics are satisfied. Although the absence of discontinuities in predetermined individual and family characteristics cannot prove the balancing property of unobservables, it is reassuring to find that individuals on both sides of the cut-off are observationally equivalent.

The inclusion of these additional individual and family controls in column 2 of Table 16 changes only modestly the estimates of the reduced-form regressions. The IV estimates at the bottom of the Table are around 20% smaller than without these controls. The moderate reduction could likely be explained by model misspecification due to the inclusion of the set of predetermined variables. (Imbens and Lemieux 2007).

## 2.7 Interpretation of the effects

A crucial question pertains to the channels through which the negative group effect operates. The substantial negative effect could either be driven by direct peer effects, e.g. through being with on average lower-performing classmates in the older class, or by

---

<sup>46</sup> Choosing different specifications for the RD by including either only a linear polynomial term or a cubic term makes the estimate for this variable insignificant, so that the single significant estimate can either be attributed to model misspecification or random chance. Any other specification for the functional form or estimating the RD without robust standard errors does not change the significance of the estimates of any of the variables.

indirect effects of the peer group composition that work through behavioural changes by students, teachers or schools to the class composition.

### **2.7.1 Exogenous peer characteristics and direct peer effects**

In the literature, it is often assumed that peer characteristics such as sex, race and socioeconomic status are proxies for (unobserved) peer ability and that exogenous peer effects work through being grouped with peers of different ability. The academic achievement of marginal students might be affected because there are more or less bright students from whom to learn or more or less students who ask stimulating questions in class.

Column 2 of Table 18 reports the estimates of the difference in mean values of a number of peer variables for students around the cut-off point. The first row reports the difference in peer age in the classrooms and the second row the difference in mean months repeated by students in the class. Unlike with the individual characteristics, I observe large and significant changes in peers' characteristics at the threshold. Peers in the older class are on average about 8 months older, which is almost completely due to the higher share of repeaters in these classes.<sup>47</sup> The remainder is due to late enrolment at first grade and temporary drop-out from school followed by re-enrolment later.

Repeaters and students who enrol late at first grade often belong to families from a more deprived socioeconomic background (Patrinos and Psacharopoulos 1996 and Gomes-Neto and Hanushek 1994), which causes the socioeconomic indicators of peers to be systematically different between the two classes. RD estimates for many of these pre-determined characteristics show a statistically significant discontinuity in peer characteristics among students around the cut-off point.

---

<sup>47</sup> Calculation based on the theoretical enrolment age of students and the number of months repeated by students show that repetition accounts for about 75% of total age-grade mismatch.

Besides mean age, age dispersion in the class also differs considerably between the two classes. With the larger number of repeaters, age dispersion in the older classes is considerably greater than in the younger classes. The standard deviation of age is 40% greater (3.5 months) in the older classes (Table 19, row 1). Figures 4 and 5 show the distribution of age of students for the two classes and give a graphical representation of the difference in the distribution of age between the classes.

Overall, students to the right of the cut-off point, while not being different from students just to the left on a whole range of individual and parental characteristics, have peer groups that not only consist of fewer girls, a higher fraction of blacks, a lower fraction of mixed students, and a higher share of children from more deprived socioeconomic background but also, due to widespread grade repetition, more heterogeneous classmates.

### **2.7.2 Indirect effects: responses of schools**

Another concern for the estimation of class peer effects is, that correlated effects in the form of common shocks to the peer group (whether exogenous or endogenous) may bias the peer effect estimates. Although it is not possible to completely rule out the existence of any differences in the learning environments between the younger and older classes, I can nonetheless assess whether there exist observable differences in a broad set of teacher and class characteristics, potentially in response to differences in the class composition.

Systematically different learning environments may arise from assigning teachers of different quality to either of the two classes. This may happen in a compensatory fashion, such that better teachers are allocated to weaker classes, which would lead to an underestimation of the peer effect. Better educated or more experienced teachers could also be allocated to the younger class to strengthen good students further, which would



lead to overestimating the peer effect. Headmasters are asked in the background questionnaire how they generally allocate teachers to classes. The vast majority (68%) of headmasters report allocating teachers in a non-systematic fashion to classes, either by means of a draw or by no specific criteria. Less than 2% of headmasters allocate more experienced teachers to stronger classes, and around 16% allocate the more experienced teachers to weaker classes. The remainder (13%) allows teachers to select the classes among themselves.<sup>48</sup>

To test whether there still are any systematic differences in teacher characteristics between the younger and older classes, I estimate teacher characteristics for the RD sample of students using the same specification as for the main estimates and the results are reported in Table 19. None of the teacher's characteristics, including sex, age, race, experience, education, training and earnings, reveal any significant difference between the two classes and the estimated coefficients are generally very small. This confirms that there is no evidence for strategic allocation of teachers. Including teacher characteristics as controls in the RD estimates (Table 16, column 3) also does not change the estimate for the peer effect in any relevant way.

Additional information from the teacher questionnaire about the allocation of teaching resources within the school to classes also provides some additional evidence that the estimates are not biased by common effects. Teachers report on the frequency of parent-teacher conferences, the quality of textbooks, and whether the provision of financial and pedagogic resources or of teaching support staff for class teaching is insufficient. None of the variables on teacher characteristics or teaching resources in the classroom reported in Table 19 are significantly different between the two groups.

---

<sup>48</sup> Unlike in settings in which teacher wages are a function of test scores, teacher wages and promotion in public schools in Minas Gerais state are mostly determined by qualification and seniority so that there is less of an economic incentive to teach better classes. Details can be found in law No. 15.293 *Establishing the Careers of Professionals in Basic Education* in the state of Minas Gerais.

As outlined above, there is some concern about the difference in class size between the older and younger classes. The estimate in Table 19 reveals that the number of students in the older class is on average lower (by the order of four students) compared to the younger class. As class size may have an effect on student achievement, this may potentially lead to a bias in the estimation of the peer group effect. There is some agreement in the literature that smaller classes may be beneficial (see Angrist and Lavy 1999 and Urquiola 2006). The effects reported in the literature are nevertheless relatively small and mostly refer to a substantial reduction in the number of students per class. In the present case, the older class is on average smaller, so that - if anything - this may lead to a downward bias of the true peer group effect on student outcomes. As the difference in class size is rather limited, it is unlikely that it leads to any considerable bias of the estimates.

### **2.7.3 Indirect effects: responses of teachers and students**

Despite the fact that teachers are observationally equivalent across classes, their teaching practices may differ as a consequence of teaching classes with a different composition of students. I use information from the student questionnaire in which students report on items related to teaching practices and the behaviour of their peer students in class. The item responses that express levels of agreement with different statements on peer and teacher behaviour, ranging from 0 to 1, have been aggregated by averaging across all the standardized outcomes at the class level.<sup>49</sup> Table 20 reports the RD estimates using the aggregated variables and the specifications as for the estimates in Table 19.

---

<sup>49</sup> Because of missing values in item responses in this part of the student questionnaire, the RD estimates on individual values are less precise. As the response of teachers is not limited to the pivotal students at the threshold, but should equally affect the other students, using the aggregated data also seems sensible, despite the potential for compositional effects. Because the coefficients from the RD on individual data are very similar to the estimates in Table 18, compositional effects nonetheless do not seem to play a relevant role in this case.

The estimates reveal that students report significantly different teaching practices across classes. This is particularly remarkable as there are no differences in observable teacher characteristics.

Students in the older class report less often that their teacher is available to clarify doubts about the course content. The coefficient is -0.04 and statistically significant at the 1% level.<sup>50</sup> Similarly, students in the older class feel that the opportunity to express their opinion in class is substantially lower (-0.029, which is about 0.3 of a standard deviation of the mean). Further evidence of an effect on teaching practices through the impact on the distribution of instruction time is given by the difference in the answers on whether the class teacher helps some students more than others. The estimate for this variable shows a 0.084 difference between classes. Teachers in the older class are compelled to distribute their attention and instructional time more unequally, possibly devoting relatively more time to specific groups of students or addressing the same material targeted at different skills levels. With a more heterogeneous group, teachers may be less able to teach to the median students, as they need to specifically address the needs of students at the tails of the distribution. The distributional features of the class composition also result in teachers being less able to devote enough time until every student has comprehended the material (-0.027).

The higher dispersion in age and ability presumably demands that teachers address different skill levels separately. In support of this hypothesis, the proportion of the planned curriculum actually taught during the school year as reported by the teacher is about 6% lower for the older classes (Table 20).

In addition to the above findings on the differences in teaching practices, information from the student questionnaire also reveals significant differences in the

---

<sup>50</sup> Given the categorical nature of the answers to these questions, interpretation is not straightforward. To give an idea about the size of the effect, the point estimate is 0.49 of a standard deviation of the mean of the variable.

behaviour of students. Students in older classes report more often that their classmates are noisy and disruptive (0.038).<sup>51</sup> With a more heterogeneous student composition teachers may need to spend more time on particular groups of students and more idle time for the remainder of students may also result in more disruptive behaviour.<sup>52</sup>

The probability of students leaving class early is also substantially higher in the older classes (0.070), which may also contribute to disruption of teaching in these classes. The less favourable teaching environment is also confirmed by students in the older class reporting more often that their teacher needs to wait to start teaching at the beginning of class due to noise (0.053).

The less favourable teaching environment may also have an effect on teacher motivation. Students of the older class report more often (0.041) that a teacher has been absent from school. The effect on absence of teachers may be interpreted as a response to the more deprived and demanding teaching environment. In turn, although difficult to quantify in terms of hours of instruction lost, teacher absence may also impact on the achievement of students, creating negative feedback effects between class composition, teacher and student behaviour.

Similarly to the findings of Lavy, Paserman and Schlosser (2012) the above results suggest that teaching practices respond to the group composition and may be an important channel in explaining the negative peer effect for students close to the class threshold.

Table 19 also shows that the percentage of students who do not participate in the PROEB test, due to illness or other reasons, differs between the two classes. Although the non-response rate differs between younger and older classes for the peer group and

---

<sup>51</sup> The difference in class behaviour reported by students is confirmed by information from the teacher questionnaire. Teachers in the older classes are more likely to report disciplinary problems with their students (0.25) (Table 17).

<sup>52</sup> Interestingly, students from the entire age rank in the older class, not only marginal students close to the threshold, report a higher level of noise and disruption, which suggests that behavioural changes are not only due to the higher share of repeaters.

is about 9% higher in the older classes, the non-response rate has a smooth transition across the discontinuity point. The size of the RD estimate for the non-participation rate at the threshold is very small and not statistically significant, so that the estimates cannot be confounded by differential non-response rate of students on either side of the cut-off point.<sup>53</sup>

#### **2.7.4 Opening the black-box of the peer-group effect: heterogeneous treatment across schools**

The previous sections have discussed different potential channels through which the peer composition in this setting may lead to the drop in academic performance of students close to the cut-off point. It remains a challenging task to distinguish the precise role of the different characteristics of peers that lead to such a large disadvantage among students in older classes, whether directly or indirectly through behavioural adjustments.

The unique setup at hand with discontinuities in more than 350 schools allows – under some assumptions – the examination of the role of different observable characteristics of the peer group in explaining the gap in academic achievement. More precisely, the fact that the difference in the characteristics of peers between children in younger and older classes differs across schools can be used to gain an understanding of the role of the different channels. For students around the cut-off point, class characteristics, such as the socioeconomic composition of their peer group, are arguably quasi-random and the difference of these characteristics between classes varies across schools and can be related to the size of the test score difference across classes at the threshold.

---

<sup>53</sup> The data appendix provides information on how the non-response rate on the class level and around the threshold has been established.

I use a two-stage minimum-distance estimator that can be easily implemented using standard statistical packages.<sup>54</sup> In the first stage I estimate the size of the discontinuity in test scores at the cut-off and the differences in peer characteristics between the two classes by 2SLS separately for each school. In the second stage, the estimated discontinuities in test scores are used as dependent variable and are regressed on the estimated differences in class characteristics  $z_{cs}$

$$b_s = \alpha_0 + \alpha_1 \Delta z_{cz} + u_s \quad (5)$$

where  $b_s$  are the estimated discontinuities in test scores for marginal students from the first stage.

Because the estimates of  $b_s$  are based on regressions using individual data, the minimum distance estimator is derived by minimizing the weighted difference between the auxiliary parameters from the first stage estimation, where the weights are equal to the reciprocal of the square of the standard errors of the first stage running minimum-distance weighted least squares. Given the quasi-random allocation of students close to the threshold, the set of class characteristics are exogenous for the marginal students at each school, so that the variation in these differences across schools can be related causally to the size of the discontinuity in test scores. Under the assumption of homogenous treatment effects, by instrumenting all class characteristics by the probability of being on either side of the cut-off point, this procedure should deliver estimates of the effect of the class characteristics that are purged of the bias induced by non-random sorting of pupils across classes by a minimum-distance RD design.

Obviously, to the extent that there are other unobservable class level characteristics that affect outcomes and are correlated with the included regressors, the minimum distance estimates will still confound the effect of such variables with the effect of the included regressors. For example, if being older is also associated with

---

<sup>54</sup> Wolfowitz (1957) introduced the minimum-distance estimator. See Kodde et al. (1990) for details.

lower innate ability, for example because older students have previously repeated a grade, but I am unable to measure ability, the measure of the average age of peers will also pick up the effect of having less able peers. It is consequently not possible to disentangle the effect of ability heterogeneity from the effect of age heterogeneity in this context.

Table 21 reports the coefficients of the above two-stage procedure.<sup>55</sup> Most of the independent class characteristics are very imprecisely estimated and the direction of the effect is puzzling for some variables. The estimate of the difference in absolute age between the two classes on the test score gap is small and not statistically significant. Also, the coefficient for the difference in mean grades repeated by students in each class is small and not significant. Although a considerable part of the differences in the mean and the variation of age is due to the different fraction of repeaters in the two classes, this does not seem to drive the negative effect estimated earlier. I even find a small negative effect on the absolute magnitude of the estimated discontinuity in test scores, but the estimate is not statistically significant.

Other coefficients reveal a relatively unsystematic pattern: some of the differences in class characteristics are positively related to the size of the discontinuity, such as the fraction of male students, the fraction of black students, or the mean number of computers available at the homes of students, while other variables show a negative relationship, such as the fraction of white and mixed students or mean books in the students' households. None of these estimates is nevertheless statistically significant at conventional levels of significance. The coefficients for the mean number of washing machines and freezers are marginally significant at the 10% level of significance. Despite the pronounced differences in various socioeconomic peer characteristics, these do not seem to play a significant role in explaining the estimated group effect. The

---

<sup>55</sup> The dependent variable of the test score gap carries a positive sign, so that a larger positive value refers to a larger negative discontinuity in math test scores between class 1 and 2.

single significant variable for explaining heterogeneity in the size of the discontinuities across schools is the difference in the age dispersion between classes. A one month difference in the standard deviation of age explains 0.035 of a standard deviation in maths test scores.

These findings are in line with the results of Hoxby and Weingarth's (2006) on the importance of the age dispersion in the reference group on academic achievement.

## **2.8 Conclusions**

In this chapter I introduce a novel way of identifying class peer effects using an RD design that exploits the rule which assigns students of a cohort to classes according to their ranking along the age distribution. The RD design allows us to compare students who are very similar in age but occur to be assigned to classes with either younger or older students. By exploiting this rule I provide evidence for strong negative effects on maths achievement for marginal students of being in a class with older peers. I find that marginal students who are assigned to the older classes have maths test scores that are around half of a standard deviation lower than those of students assigned to the younger classes.

Concurrently with being in different peer environments, marginal students are also either the oldest or the youngest in their respective classes and, apart from the effect from being assigned to classes with different peer characteristics and their distribution, there could be a separate pure relative age effect at work. It is nevertheless debatable whether conceptually there is a difference between a potential pure relative age effect and an age peer group effect and, given the identification strategy, these effects would by definition be practically indistinguishable. Moreover, there is no



evidence for the existence of a separate pure relative age effect elsewhere in the literature.<sup>56</sup>

The results contribute to the debate on streaming or tracking of students into classes and schools and have direct policy implications. If the distribution of underlying observable student characteristics in classes has substantial effects on achievement, while changes in the composition could possibly be achieved at zero cost at the school level, there is a strong case for sorting students into classes aiming at a more homogenous class composition in terms of age or ability. Findings in the related literature point to a potential trade-off between direct and indirect peer effects from grouping students by age or ability. Zimmer (2003) finds that tracking in the US has a positive effect even on low-achieving students through more tailored instruction and can outweigh the negative direct effect on low-achievers from the absence in high quality peers. When grouping students according to ability, low-achieving students may no longer benefit from the presence of high-achieving peers, but instead may take advantage of the lower variation in ability, potentially leading to a more efficient teaching environment for all students. It is particularly important to consider this trade-off in educational systems with substantial age and ability heterogeneity, as is the case in many low- and middle-income countries.

The findings in this chapter also contribute to the understanding of policies that aim at reducing the age variation in cohorts of students. Policies designed to reduce late enrolment in primary schools may have positive effects on all students by reducing the age heterogeneity in each cohort. Correspondingly, grade retention policies may have a

---

<sup>56</sup> Using experimental data from Project STAR, Cascio and Whitmore Schanzenbach (2007) do not find evidence for an effect from relative age on mean test scores. Elder and Lubotsky (2009) also show that a commonly postulated positive relationship between achievement and school entry age is primarily driven by the skills older children acquired prior to kindergarten rather than absolute or relative age effects. As the identification strategy employed in this chapter is based on the discontinuity around the median age in the cohort, the estimated effects are not confounded by relative age effects at the extremes of the age distribution, i.e. being the youngest or oldest in the cohort, so that targeting the curriculum to a specific age group will not bias the estimated effects.

significant impact on the dispersion of age in cohorts of students and may therefore affect achievement of all students. Retained students increase the dispersion of age in the cohort and may impose a negative externality on all students in the class regardless of the existence of direct and indirect effects of having (low-achieving) repeaters in the peer group.

**Table 15: Means and Proportions of Student and Teacher Characteristics**

Panel A: Student characteristics		Younger class		Older class	
	Class size	24.738	(0.287)	21.868	(0.302)
Age	(in years)	10.930	(0.009)	11.670	(0.014)
Sex	Female	0.524	(0.005)	0.458	(0.006)
Race	White	0.306	(0.005)	0.264	(0.005)
	Mixed	0.526	(0.005)	0.517	(0.006)
	Black	0.097	(0.003)	0.143	(0.004)
	East-Asian	0.027	(0.002)	0.034	(0.002)
	Indigenous	0.044	(0.002)	0.042	(0.002)
Repeater	Never repeated	0.797	(0.004)	0.489	(0.006)
	Repeated once	0.142	(0.004)	0.292	(0.005)
	Repeated twice	0.043	(0.002)	0.148	(0.004)
	Repeated 3 or more times	0.018	(0.001)	0.070	(0.003)
SES	Family with Bolsa Família	0.480	(0.005)	0.592	(0.006)
	Household employs domestic worker	0.137	(0.004)	0.113	(0.004)
	Number of books	23.496	(0.322)	19.428	(0.330)
	Number of cars	0.608	(0.009)	0.503	(0.009)
	Number of computers	0.262	(0.005)	0.195	(0.005)
	Number of fridges	0.999	(0.005)	0.958	(0.006)
	Number of freezers	0.302	(0.006)	0.282	(0.007)
	Number of radios	1.342	(0.008)	1.286	(0.009)
	Number of TVs	1.497	(0.008)	1.396	(0.009)
	Number of DVD players	0.849	(0.007)	0.786	(0.008)
	Number of bathrooms	1.246	(0.006)	1.175	(0.006)
	Number of washing machines	0.758	(0.007)	0.752	(0.007)
	Number of tumble dryers	0.168	(0.005)	0.163	(0.005)
Panel B: Teacher characteristics					
Sex	Female	0.983	(0.011)	0.965	(0.015)
Age	(in years)	40.495	(0.468)	40.094	(0.486)
Race	White	0.456	(0.030)	0.477	(0.030)
	Mixed	0.420	(0.029)	0.399	(0.029)
	Black	0.093	(0.017)	0.081	(0.016)
	East-Asian	0.028	(0.010)	0.039	(0.012)
	Indigenous	0.004	(0.004)	0.004	(0.004)
Highest	Secondary education	0.100	(0.018)	0.118	(0.019)
edu. degree	Higher education – pedagogic degree	0.210	(0.024)	0.208	(0.024)
	Higher education - regular	0.410	(0.029)	0.389	(0.029)
	Higher education and teaching qualification	0.203	(0.024)	0.174	(0.022)
	Higher education – other	0.076	(0.016)	0.111	(0.019)
	Earnings (in R\$)	771.74	(22.803)	743.60	(23.754)
	Years of experience in education	14.023	(0.360)	13.862	(0.375)
	Participation in continuing education	0.375	(0.028)	0.363	(0.029)

Notes: The data from the upper panel are taken from the student background questionnaires, the data from the lower panel are from the teacher questionnaires. Source: PROEB 2007.

**Table 16: Main Estimation Results**

	(1)	(2)	(3)
Panel A: first stage			
Dependent variable: class rank			
	0.467***	0.453***	0.451***
	(0.056)	(0.057)	(0.056)
R <sup>2</sup>	0.326	0.370	0.403
Panel B: reduced form			
Dependent variable: maths test scores			
	-26.445***	-19.196**	-19.513**
	(7.458)	(7.646)	(7.743)
R <sup>2</sup>	0.405	0.482	0.485
Panel C: IV regression discontinuity results			
Dependent variable: maths test scores			
	-56.574***	-42.385***	-43.297***
	(15.299)	(15.455)	(15.673)
R <sup>2</sup>	0.410	0.485	0.489
Observations:	1,688	1,688	1,688
School fixed effects	yes	yes	yes
Individual controls	no	yes	yes
Teacher controls	no	no	yes

Notes: The top panel reports the first stage regressions using OLS estimating equation (4). The middle panel reports the coefficient on maths test score on the dummy equal 1 for the age rank larger than 0 (reduced form). Test scores are centred using school fixed effects in all specifications. The bottom panel reports IV estimates of the effect of being in the older class on maths test scores, where being in the older class has been instrumented by a dummy for having an age rank larger than 0. All specifications include a second-order polynomial in the age rank. Specifications in column (2) include the whole set of predetermined individual and family characteristics, including sex, race, repeated years and SES family characteristics; specifications in column (3) additionally include all predetermined teacher characteristics, including teacher sex, race, age, salary, variables on educational background and experience. Heteroskedasticity consistent standard errors are clustered by schools and reported in parenthesis. \*\* and \*\*\* denote significance at the 5% and 1% level, respectively.

**Table 17: RD Estimates of Maths Test Scores**

	Ranks from threshold in months				
	1 month	2 months	3 months	4 months	5 months
	Estimated discontinuity at threshold				
Quadratic	-56.574*** (15.299)	-54.578*** (12.561)	-59.044*** (11.103)	-57.193*** (10.791)	-59.182*** (10.653)
Cubic	-55.477*** (15.551)	-54.467*** (12.622)	-59.560*** (11.106)	-57.188*** (10.842)	-58.416*** (10.722)
Quadratic with full controls	-43.297*** (15.673)	-43.762*** (12.446)	-45.216*** (11.259)	-43.600*** (10.980)	-43.066*** (10.675)
Cubic with full controls	-41.689** (16.299)	-43.753*** (12.45)	-45.625*** (11.274)	-43.769*** (11.031)	-42.726*** (10.749)
Number of student observations	1,688	3,142	4,547	5,884	7,223

Notes: The dependent variable is the maths test score and entries are estimates of the discontinuity including the different range of observations in terms of the age rank indicated by the column heading. Entries for row (1) are the estimated coefficients of the RD from models that include a quadratic polynomial in the age rank for the different range of observations. Row (2) includes a cubic polynomial in the age rank. Rows (3) and (4) additionally include the full set of controls as in column (4) of Table 16. Heteroskedasticity consistent standard errors are reported in parentheses. \*\* and \*\*\* denote significance at the 5% and 1% level, respectively.

**Table 18: RD Estimates of Predetermined Individual and Family Variables**

		(1)		(2)	
		Individuals		Peers	
Fraction of:	Age (in months)	0.442	(0.735)	8.157***	(0.796)
	Grades repeated (in months)	0.728	(0.879)	7.487***	(0.457)
	Female	0.190	(0.127)	-0.088***	(0.019)
	White	0.008	(0.092)	-0.035	(0.023)
	Mixed	-0.037	(0.102)	-0.072**	(0.032)
	Black	0.115**	(0.055)	0.089***	(0.018)
	East-Asian	-0.026	(0.022)	0.011	(0.009)
	Indigenous	-0.076	(0.047)	-0.001	(0.009)
	Domestic helper	-0.020	(0.058)	-0.053***	(0.017)
	Bolsa Família	0.165*	(0.099)	0.144***	(0.027)
Number of:	Parental homework support	0.027	(0.054)	-0.066***	(0.016)
	Bathrooms	-0.101	(0.098)	-0.129***	(0.033)
	Books	-4.314	(4.956)	-8.016***	(1.928)
	Cars	-0.167	(0.138)	-0.141***	(0.039)
	Computers	-0.031	(0.068)	-0.108***	(0.022)
	Fridges	0.096	(0.077)	-0.074**	(0.031)
	Freezers	-0.013	(0.087)	-0.052**	(0.025)
	Radios	0.195	(0.158)	-0.083	(0.052)
	Washing machines	0.080	(0.105)	-0.037	(0.033)
	Dryers	-0.057	(0.082)	0.014	(0.021)
	DVDs	0.125	(0.121)	-0.120***	(0.035)
	TV sets	-0.008	(0.141)	-0.194***	(0.042)
	Video players	0.080	(0.107)	-0.066**	(0.028)
Number of student observations		1,688		1,688	

Notes: Entries are separate IV estimates of the class effect on student and family characteristics, where being in the second class has been instrumented by a dummy for having an age rank larger than 0. For each variable a separate regression has been estimated. Column (1) reports the effect around the discontinuity point for the individual values of the characteristics; column (2) reports the estimates for the values of the peer group characteristics for the same individuals around the cut-off point. All specifications include a second-order polynomial in the age rank Heteroskedasticity consistent standard errors, clustered on the school level are reported in parentheses. \*, \*\* and \*\*\* denote significance at the 10%, 5% and 1% level, respectively.

**Table 19: Class and Teacher Characteristics**

Dependent variable			
Class characteristics	Std. deviation of age (in months)	4.012***	(0.381)
	Class size	-4.162***	(0.583)
	Non-participation rate (at threshold)	0.006	(0.004)
	Non-participation rate (of peers)	0.093***	(0.022)
Teacher characteristics	Female	-0.087*	(0.049)
	Age (in years)	-1.607	(1.615)
	White	-0.005	(0.101)
	Mixed	-0.048	(0.103)
	Black	0.025	(0.060)
	East-Asian	0.020	(0.033)
	Indigenous	0.009	(0.009)
	Higher education degree	0.030	(0.077)
	Postgraduate degree	-0.034	(0.103)
	Years passed since graduation	-0.108	(0.226)
	Earnings (in Brazilian Reais)	-69.176	(56.943)
	Participation in continuing education	-0.015	(0.091)
	Experience in education (in years)	-0.395	(0.259)
	Teacher has other source of income	-0.089	(0.093)
Teaching resources	Frequency of parent-teacher conferences	0.068	(0.135)
	Quality of textbooks	0.178	(0.098)
	Insufficient financial resources	-0.024	(0.080)
	Insufficient pedagogic resources	-0.063	(0.108)
	Insufficient teaching support staff	0.036	(0.102)
Number of student observations		1,688	

Notes: Entries are separate IV estimates of the class effect on class and teacher characteristics, where being in the second class has been instrumented by a dummy for having an age rank larger than 0. For each variable a separate regression has been estimated. The data come from the teacher questionnaire of PROEB 2007 and the School Census (for class characteristics). Class teacher statements come from the teacher questionnaire and relate to the specific class taught. Class size is calculated using the official number of students enrolled in a class based on information from the School Census. The *non-participation rate (at threshold)* is based on the difference in the distribution of students of age ranks between the school census and PROEB test takers. The *non-participation rate of peers* is based on the difference between class size and number of students participating in the PROEB test. The variable *quality of textbooks* ranges between 0 and 1, with the value 1 given for the best quality and 0 for the lowest. All regressions control for school fixed effects. Heteroskedasticity consistent standard errors are reported in parentheses. \* and \*\*\* denote significance at the 10% and 1% level, respectively.

**Table 20: Response of Teaching Practices to Class Composition**

Disciplinary problems with students	0.253**	(0.113)
Rate of planned curriculum taught	-0.059***	(0.019)
Rate of students expected to finish primary school	-0.097***	(0.023)
Rate of students expected to finish secondary school	-0.133***	(0.031)
Teacher availability to clarify doubts	-0.039***	(0.008)
Teacher explains until all students understand	-0.027***	(0.009)
Teacher gives opportunity to express oneself	-0.029***	(0.010)
Teacher helps more some students	0.084***	(0.015)
Teacher interested in learning progress	-0.028***	(0.007)
Teacher needs to wait to start teaching	0.053***	(0.017)
Teacher absenteeism	0.041***	(0.012)
Fellow students leave classroom early	0.070***	(0.015)
Fellow students are noisy and disruptive	0.038***	(0.015)
Fellow students learn taught material	-0.044***	(0.010)
Fellow students pay attention in class	-0.009	(0.010)
Teacher enforces student attention	-0.010	(0.007)
Teacher corrects homework	-0.020	(0.013)
Number of student observations	1,688	

Notes: Entries are separate IV estimates of the class effect on the response of teachers and students to the class composition, where being in the older class has been instrumented by a dummy for having an age rank larger than 0. For each variable a separate regression has been estimated. The variables in the top panel are from the teacher questionnaire. The variable *disciplinary problems with students* is a dummy taking a value 1 if teachers report that there are problems with the discipline of students. The variables from the bottom two panels come from the student questionnaire of PROEB 2007. The variables have been recoded from categories ranging from “totally disagree” to “totally agree” on a scale from 0-1 and aggregated on the class level. All regressions control for school fixed effects. Heteroskedasticity consistent standard errors, clustered on the school level, are reported in parentheses. \*\* and \*\*\* denote significance at the 5% and 1% level, respectively.

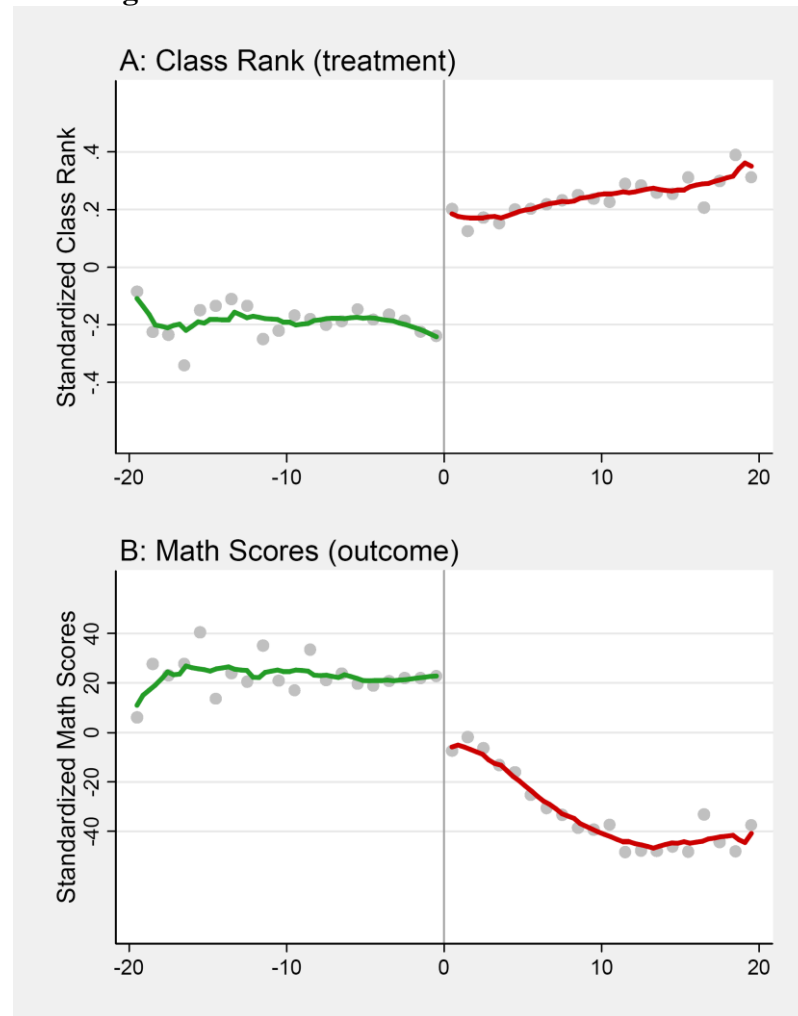


**Table 21: Heterogeneous Treatment across Schools**

Difference in class means		
Age dispersion	3.485**	(1.446)
Mean age (in months)	-0.704	(1.625)
Mean grades repeated (in months)	-2.149	(23.899)
Fraction of male students	26.985	(26.070)
Fraction of white students	-21.781	(35.988)
Fraction of mixed students	-35.811	(26.031)
Fraction of black students	23.887	(50.446)
Fraction of Asian students	-81.925	(106.722)
Fraction of households with domestic workers	-1.419	(49.267)
Fraction of households receiving Bolsa Família	-29.259	(35.020)
Mean books	-30.246	(18.638)
Mean bathrooms	-2.999	(32.698)
Mean cars	0.302	(25.877)
Mean computers	4.398	(43.518)
Mean fridges	-16.989	(28.483)
Mean freezers	-58.830*	(33.266)
Mean radios	31.154	(24.689)
Mean washing machines	41.407*	(22.327)
Mean DVD players	42.635	(32.030)
Mean TV sets	-3.199	(24.698)
Mean video players	34.795	(33.711)
Teacher controls	yes	
Number of observations (discontinuities):	363	
R <sup>2</sup>	0.302	

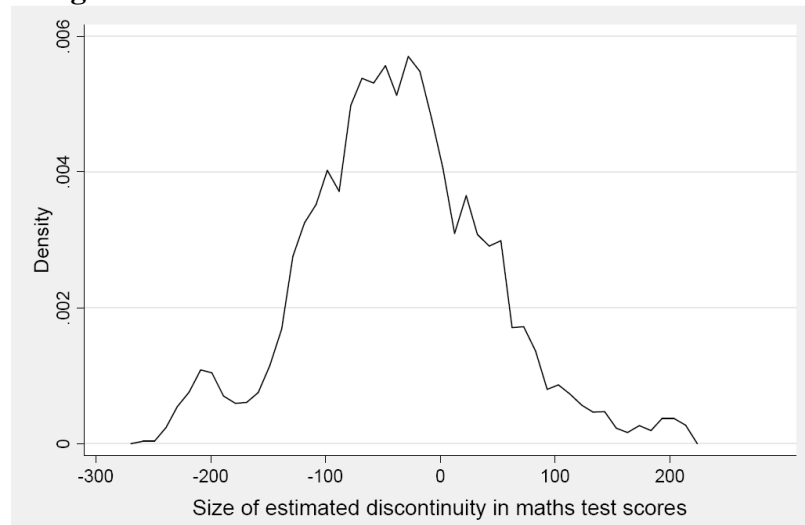
Notes: The dependent variables are measures of the absolute size of the discontinuities in math test scores at the cut-off point on the school level estimated by 2SLS. The entries report coefficients from the second stage of the minimum distance estimation, where the weights are equal to the inverse of the standard errors of the estimates of the first stage. Independent variables are the estimated differences in means of the peer values of socioeconomic characteristics, class age and its distribution. Heteroskedasticity robust standard errors are reported in parenthesis. \* and \*\* denote significance at the 10% and 5 level, respectively.

**Figure 3: Local Averages and Local Linear Regression of Treatment and Outcome Variable**



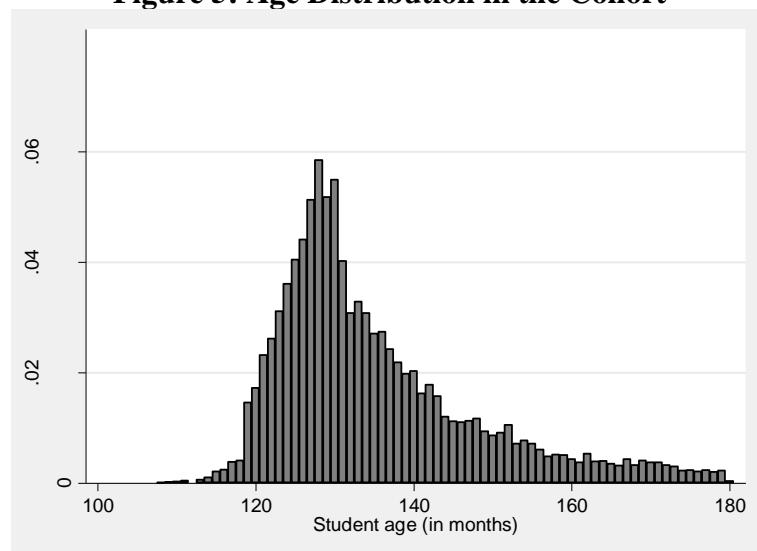
Notes: The graphs plot local averages of the standardized class rank of students and of the standardized maths test score according to the age ranking in the cohort as distance of students from the cut-off point and local linear regression fits on both sides of the cut-off point using a rectangular kernel with a bandwidth of 3 months.

**Figure 4: Distribution of RD Estimates Across Schools**



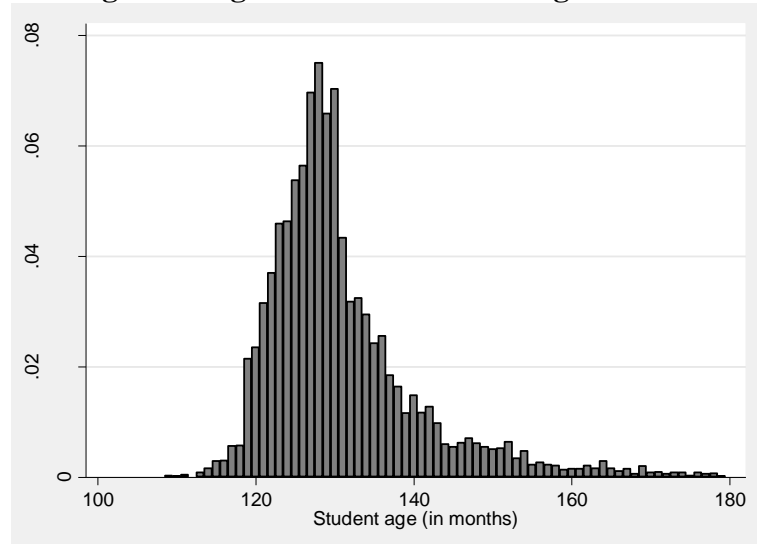
Notes: The graph plots kernel density estimates of school specific estimated discontinuities using a rectangle kernel with a bandwidth of 20.

**Figure 5: Age Distribution in the Cohort**



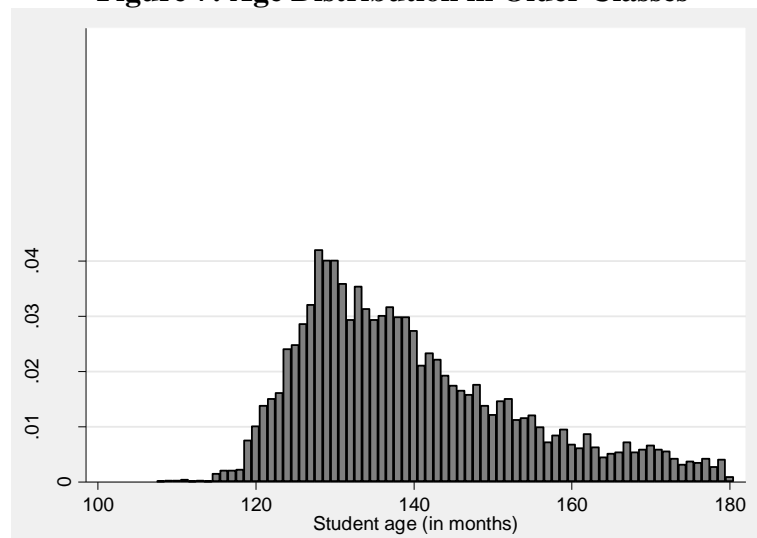
Notes: The graph plots the density of student age for all students in the cohort, age is reported in months.

**Figure 6: Age Distribution in Younger Classes**



Notes: The graph plots the density of student age for class 1 (younger class), age is reported in months.

**Figure 7: Age Distribution in Older Classes**



Notes: The graph plots the density of student age for class 2 (older class), age is reported in months.

## 2.9 Data appendix

This appendix describes the variables of students, teachers, class, headmaster and schools used in this chapter.

### *Outcome variable: Maths test score*

The PROEB test score for mathematics has been constructed from a battery of 40 multiple choice questions covering four areas: space and shapes, size and measurement, algebraic operations, and treatment of information. For each question, students are offered 4 possible answers, of which one is correct. The test scores have been standardized to a mean of 500 and a standard deviation of 100. The test is administered in November, close to the end of the school year.

### *Student socioeconomic characteristics*

All information on the socioeconomic background of students comes from a socioeconomic questionnaire which is a supplement to the maths test. *Racial affiliation* is self-reported by students, as well as all other information on the background characteristics of the students and their families.

The dummy variable *Bolsa Família* reports whether the family is a recipient of cash-transfers from the federal programme and takes a value of 1 if the family is a recipient.

The dummy variable *domestic worker* records whether the family employs one or more domestic workers (part-/full-time).

The variables on the *number of books, cars, computers, fridges, freezers, radios, TVs, DVD players, bathrooms, washing machines and tumble dryers* are numeric and can take the values “0”, “1”, “2” or “3 and more”. The value of “3 and more” has been coded with a value of 3.

The variable *individual age* of students has been created based on three questions related to age. Students need to provide their age in years, their month of birth and indicate whether or not they have already passed their birthday in the current calendar year. This information, together with the test date of PROEB, allows the age of the children in years and months to be established. The average age of students is 135.28 months, which is approximately 11.27 years. This is about 9 months above the appropriate age at the end of 5<sup>th</sup> grade. Average age in the younger classes is 131 months and in the older class 140 months. The standard deviation of age in the cohort at 5<sup>th</sup> grade is 12.09 months. The distribution of age in the two classes differs quite considerably with a standard deviation of age in the younger classes of 10.02 months and of 14.16 months in the older classes. Figure 5 plots the density of age in the entire cohort and shows how the age distribution is skewed to the right. The histograms of Figures 4 and 5 show the different distribution of age in the two classes. Both distributions are positively skewed, with the mass of the distribution concentrated to the left. This is due because age is censored at the left tail with a minimum enrolment age of 5½, and the upper age limit. The maximum observed age is 15 years, which is almost 4 years above the average age and 5½ years above the possible youngest age. The substantial age-grade distortion in the student cohort can mostly be attributed to grade repetition by students. Every year repeated by a student contributes to the age variation based on the distribution of birth dates and the enrolment cut-off point at 1<sup>st</sup> grade. With 20% of students having repeated one year, 9% having repeated twice and 4% having repeated three or more times, repetition accounts almost wholly for the age-grade distortion observed in the data (grade repetition accounts for approximately half a year in mean student age). The remainder is likely to be due to some late enrolment and school dropout with re-enrolment or change of school by students who are then reassigned to a lower grade. Unfortunately, I do not have available information on

enrolment age for the cohort of interest. From the School Census 2007 that contains information on age for individual students for 1<sup>st</sup> grade, I can calculate that late enrolment is responsible for about 1.8 months, which is likely to be similar to the effect of late enrolment in the cohort of consideration that had enrolled 4 years earlier.

*Teacher characteristics, statements of teachers on class teaching environment and class characteristics*

The information on teacher characteristics comes from two sources and these are matched by school and class identifiers and the subject the teachers teach. All information on socioeconomic characteristics (in panel B of Table 15) comes from the annual Brazilian School Census that collects information on school, teacher and headmaster characteristics from all Brazilian schools. The variables *years passed since graduation* and the different variables on teacher professional experience have been transformed using midpoints of the ranges reported in the questionnaire. Salary of teachers is reported in Brazilian Reais (R\$1 was worth approximately US\$0.58, as of 10<sup>th</sup> September 2010) and is calculated from the mid-points of the salary ranges given in the questionnaire.

The information on the teaching environment and student behaviour comes from the background questionnaire of PROEB that is completed by all teachers. *Frequency of class council meetings* is reported as *never, once, twice and three times and more*, the last of which has been recoded with a value of 3, and reports on the adequacy of financial and pedagogic resources for class teaching are dummy variables taking a value of 1 if teachers think that resources are insufficient and 0 otherwise. Teacher statements about the progress of teaching and students have been included from the Prova Brasil 2007 teacher background questionnaire. The percentages of the *planned curriculum*

*taught*, and students finishing primary and secondary school, have been calculated using the midpoint of the percentage ranges reported by teachers.

The variable on standard deviation of age in the classes is calculated using individual student age. The variable class size is based on information from the official School Census that reports the number of students in each class.

The *Non-participation rate* at the class level is based on the difference between official student numbers as recorded by the School Census and the number of students participating in the PROEB test at the class level. The *Non-participation rate at the threshold* is established using information from the school census on the complete age distribution of all students at school. The difference between official numbers and numbers of students taking the PROEB test for the same age rank (as in monthly intervals) at the school level informs about the missing students and the non-response rate of students for each age rank at schools.

#### *Student evaluation of teaching practices and classroom environment*

The information about teaching practices and the classroom environment in Table 20 come from the student background questionnaire of PROEB. The variables report means on the class level. The bottom four variables refer directly to classmates of students, whereas the top variables refer to teaching practices and teacher behaviour. Students report their level of agreement to statements about teaching practices and the behaviour of their classmates, from strongly disagree to strongly agree. The categories have been recoded to range from 0 (strong disagreement) to 1 (strong agreement). Table 24 reports the mean and standard deviation for these variables on the class level.



### *School characteristics and headmaster characteristics*

The information on physical school characteristics comes from the annual School Census.

The dummy variable *urban school* takes a value of 1 for being in an urban setting, and 0 for a rural setting.

The dummy variable *state school* takes a value of 1 for the school being under the direct administration of the state secretariat of education in Minas Gerais and 0 for a municipal school which is under administration of the municipal secretariat of education.

The variables of *headmaster office*, *faculty room*, *school library*, *video facilities*, *TV room*, *copy machine*, *printer*, *overhead projector*, *school kitchen*, *internet access*, *computer and science lab*, *filtered water*, *public water supply*, *public sewerage* and *sport facilities* are all dummy variables taking a value of 1 when the facilities exist at the school and 0 otherwise.

### *Normalization on school and class level*

As mentioned in the text, each of the regressions includes school fixed effects. For this purpose all variables used for the RD analysis have been normalized to have a mean of 0 at school level. Furthermore, the ranking of students has been centred on a cut-off point of 0, reporting the age rank as distance from the cut-off point.

## **2.10 Appendix on initial class assignment and class transition**

Primary education in Brazil is divided into two stages. The first stage (*initial years*) comprises five years and the second stage (*final years*) the remaining four years of primary education. During the initial years a single class teacher (*professor regente*) teaches the entire curriculum covering all subjects (mathematics, Portuguese, science,

history, geography), whereas classes are taught by specialized teachers separately for different subjects in the last four years of primary school.<sup>57</sup> The aim of the initial years, besides the achievement of curriculum targets, is to establish social and emotional ties and to build the capacity of students in interacting with other children of similar age and with adults.<sup>58</sup> To facilitate this aim, all subjects are taught by a single class teacher and students remain in their originally assigned class formed at first grade throughout the first five years of primary school. It is therefore informative to learn about the initial assignment of students into classes and the transition of students from grade to another. As PROEB only focuses on the cohorts tested (5<sup>th</sup> and 9<sup>th</sup> grade of primary school) there is nevertheless no individual data for the initial class assignment at first grade for the cohort of interest. With a change in the data collection method of the Brazilian school census in 2007, information on individual students rather than aggregated class and school data is collected from the year 2007. The school census contains information on individual characteristics on age, sex and the racial attribution of students and permits to test whether or not the balancing properties of these predetermined characteristics are satisfied for the entry cohort of 2007. Table 25 reports the RD estimates for these characteristics for the 2007 entry cohort of first graders. The coefficients of the RD estimates for sex or the racial attributes of students at the threshold are relatively small and not statistically significant, confirming that the predetermined characteristics are balanced across the threshold for first grade students of the entry of cohort of primary school.

Using the 2007 and 2008 census I can follow the cohorts of students from school year 2007 to school year 2008.<sup>59</sup> The transition information includes the records of students being promoted from 1<sup>st</sup> to 2<sup>nd</sup>, from 2<sup>nd</sup> to 3<sup>rd</sup>, from 3<sup>rd</sup> to 4<sup>th</sup> and from 4<sup>th</sup> to 5<sup>th</sup>

---

<sup>57</sup> Details are outlined in Resolution SEE No 1086 of the State Secretariat of Education Minas Gerais.

<sup>58</sup> Brazilian Ministry of Education (2004).

<sup>59</sup> For this exercise the student information is available only for a restricted sample of schools. Not all schools have recorded consistently the student information across years in the school census, so that the information on the transition of classes is only available for 55 schools.

grade and I have pooled all the cohorts together. Over 80% of students remain in the class with the same peers. Conditional on regular transition, 96% of students remain in the same peer environment. Regressing the probability of being in class  $j$  ( $j=1/2$ ) at time  $t-1$  on the probability of being in group  $j$  at time  $t$ , conditional on age rank at age  $t-1$  does not reveal a significant difference for remaining with the same class for students that rank close to the cut-off point, so that students close to the threshold have the same (high) probability to stay with the same class the next year compared to students further away from the threshold.

## **2.11 Appendix on selection of schools in the sample**

In this chapter I use schools with two classes in a given cohort only; schools with a single class are excluded given the RD identification strategy. I excluded schools with more than two classes for two reasons: This precludes to install ‘special’ classes of low performing students, students with behavioural problems, or students with other specific observable or unobservable characteristics, such as repeaters. As these may be removed strategically from the cut-off this could invalidate the assumption of random assignment around the class cap. This also ensures that there is sufficient variation in class characteristics to estimate meaningful peer effects.

It may nevertheless be interesting to understand whether students used in this paper are different from students in all the other schools. In Table 23 I report the socio-economic composition of students in schools with 2 classes with students from all other schools. Students look very similar in terms of their socio-economic composition. The fraction of white and black students in the schools used in this chapter is slightly smaller, in exchange for more mixed students. Overall students used in the analysis seem to have a slightly better socio-economic status, as evidenced by the smaller fraction of student’s receiving Bolsa Família and the larger number of books present in

the household. In the contrary though, students from schools with two classes have on average less cars available compared to the students from all other schools; there is essentially no difference in the fraction of families employing a domestic worker. Overall, the socio-economic composition of students at schools with two classes does not differ largely from students at all other schools.

**Table 22: Means and Proportions of School and Headmaster Characteristics**

Physical school characteristics			
Means	Permanent class rooms	10.250	(0.190)
	Number of total staff	46.110	(1.150)
	Class size	23.199	(0.217)
Proportions	Urban school	0.910	(0.020)
	State school	0.553	(0.030)
	Municipal school	0.447	(0.030)
	Headmaster office	0.897	(0.016)
	Faculty room	0.844	(0.019)
	School library	0.825	(0.020)
	Video facilities	0.356	(0.010)
	TV room	0.979	(0.007)
	Video player	0.902	(0.015)
	DVD player	0.847	(0.019)
	Copy machine	0.370	(0.025)
	Printer	0.903	(0.017)
	Overhead projector	0.788	(0.023)
	School kitchen	0.926	(0.013)
	Internet connectivity	0.589	(0.028)
	Computer laboratory	0.355	(0.025)
	Science laboratory	0.106	(0.016)
	Facilities for disabled children	0.820	(0.020)
	Filtered water	0.989	(0.005)
	Public water supply	0.950	(0.011)
	Public energy supply	0.997	(0.003)
	Public sewerage	0.828	(0.019)
	Waste collection	0.913	(0.015)
	Sport facilities	0.606	(0.027)
Headmaster characteristics			
Sex	Female	0.860	(0.020)
Race	White	0.452	(0.028)
	Mixed	0.427	(0.028)
	Black	0.068	(0.014)
	Asian	0.046	(0.014)
	Indigenous	0.007	(0.005)
	Age (in years)	43.100	(0.054)
Highest educational level	Secondary education	0.050	(0.123)
	Higher education – pedagogic degree	0.318	(0.026)
	Higher education – maths	0.428	(0.028)
	Higher education – literature	0.053	(0.013)
	Higher education – other	0.151	(0.020)
	Earnings (in R\$)	1635.49	(38.85)
	Years of experience in education	18.090	(0.209)
	Years of experience at this school	6.210	(0.241)
	Years of experience as headmaster	6.949	(0.258)
	Participation in continuing education	0.114	(0.020)

Notes: Data for the physical school characteristics comes from the annual Brazilian School Census, headmaster characteristics come from the 2007 wave of PROEB.

**Table 23: Socioeconomic Characteristics of Students in Schools in Sample/ not in Sample**

Schools	In sample		Not in sample	
	(1)		(2)	
Age (in years)	10.789	(0.009)	10.716	(0.002)
Fraction male	0.492	(0.004)	0.493	(0.001)
Fraction white	0.286	(0.004)	0.303	(0.001)
Fraction mixed	0.521	(0.004)	0.497	(0.001)
Fraction black	0.119	(0.003)	0.123	(0.001)
Fraction Asian	0.031	(0.001)	0.033	(0.000)
Fraction Indigeneous	0.043	(0.002)	0.044	(0.000)
Family with Bolsa Família	0.467	(0.004)	0.498	(0.001)
Family employs domestic worker	0.127	(0.003)	0.124	(0.001)
Number of books	22.020	(0.224)	21.415	(0.054)
Number of cars	0.560	(0.006)	0.619	(0.002)

Notes: The table reports socio-economic characteristics of students in schools in the sample used for the analysis (2 classes in 5<sup>th</sup> grade) and for students not in the sample (one class, or three and more classes in 5<sup>th</sup> grade. Colum (1) reports mean characteristics of students ion schools with two classes per cohort, and column (2) characteristics of students in any other school. Data comes from the socio-economic questionnaire of the 2007 wave of PROEB.

**Table 24: Choice of Class Assignment Rule**

	coefficient	s.e.
<b>SCHOOL PHYSICAL CHARACTERISTICS</b>		
Urban school	0.025	(0.121)
State school	0.006	(0.06)
Number of permanent class rooms	0.01	(0.01)
Total number of staff	-0.001	(0.002)
Size of cohort	-0.008 ***	(0.003)
School library	0.029	(0.081)
Headmaster office	-0.224 ***	(0.067)
Faculty room	0.05	(0.083)
Video facilities	0.038	(0.091)
TV room	-0.133	(0.189)
Copy machine	0.078	(0.051)
Printer	-0.079	(0.086)
Overhead projector	-0.009	(0.07)
School kitchen	0.069	(0.081)
Internet access	0.085 *	(0.05)
Computer lab	-0.111 *	(0.056)
Science lab	-0.041	(0.09)
Filtered water	-0.019	(0.101)
Public water supply	0.145	(0.252)
Public sewerage	0.022	(0.08)
Sport facilities	-0.007	(0.048)

Table 24 continued

HEADMASTER CHARACTERISTICS			
Highest education obtained	Male	-0.019	(0.060)
	Age	-0.001	(0.002)
	Mixed	0.066	(0.054)
	Black	0.052	(0.080)
	Asian	0.167 **	(0.078)
	Indigenous	0.036 **	(0.144)
	High school	0.117	(0.188)
	Higher education - pedagogic degree	0.014	(0.129)
	Higher education – normal	-0.077	(0.126)
	Higher education & teaching qualification	0.033	(0.152)
	Higher education – other	0.044	(0.132)
	Experience in years as headmaster	-0.001	(0.002)
	Experience in years in education	0.000	(0.000)
	Continuing education	-0.055	(0.0760)
	Earnings	0.000	(0.000)
MEAN TEACHER CHARACTERISTICS			
	Proportion male	-0.113	(0.180)
	High school	-0.149	(0.116)
	Higher education – pedagogic degree	-0.056	(0.104)
	Higher education – regular	-0.081	(0.107)
	Higher education and teaching qualification	-0.128	(0.153)
	Higher education – other	-0.126	(0.126)
	Earnings	0.000	(0.000)
	Mean experience in education	0.004	(0.005)



Table 24 continued

Proportion Bolsa Família	0.147	(0.227)
Mean books	0.008	(0.006)
Proportion female	0.000	(0.275)
Mean HH with domestic worker	0.260	(0.471)
Proportion white	-3.325	(1.983)
Proportion mixed	-2.365	(1.951)
Proportion black	-2.856	(1.830)
Proportion Asian	-4.507 *	(2.294)
Proportion Indigenous	-2.984	(1.853)
Mean automobiles	-0.258	(0.176)
Mean computers	0.091	(0.298)
Mean fridges	-0.539 **	(0.259)
Mean freezers	0.381	(0.273)
Mean radios	0.050	(0.153)
Mean washing machines	0.096	(0.154)
Mean tumble dryer	0.144	(0.403)
Mean DVD players	0.109	(0.200)
Mean TV sets	-0.013	(0.155)
Mean bathrooms	0.317	(0.229)
Mean videos	-0.264	(0.250)
Constant	4.157 *	(1.878)
Observations	363	
R-squared	0.236	

Notes: The coefficients come from a linear probability model on the selected assignment rule of students into classes, where the outcome is a dummy taking a value of 1 if students are assigned to classes using their relative age to form *homogenous* classes, and 0 otherwise. Heteroskedasticity robust standard errors are reported in parenthesis. \*, \*\* and \*\*\* denote significance at the 10%, 5% and 1% level, respectively.

**Table 25: Means of Student Statements on Teaching Practices and Peer Behaviour**

	Mean	Std. Dev.
Teacher enforces student attention	0.918	(0.067)
Teacher corrects homework	0.788	(0.119)
Teacher availability to clarify doubts	0.904	(0.079)
Teacher explains until all students understand	0.891	(0.083)
Teacher gives opportunity to express oneself	0.850	(0.098)
Teacher helps more some students	0.254	(0.147)
Teacher interested in learning progress	0.917	(0.072)
Teacher needs to wait to start teaching	0.581	(0.166)
Teacher absenteeism	0.237	(0.130)
Fellow students leave classroom early	0.269	(0.150)
Fellow students are noisy and disruptive	0.527	(0.140)
Fellow students learn taught material	0.866	(0.087)
Fellow students pay attention in class	0.652	(0.108)

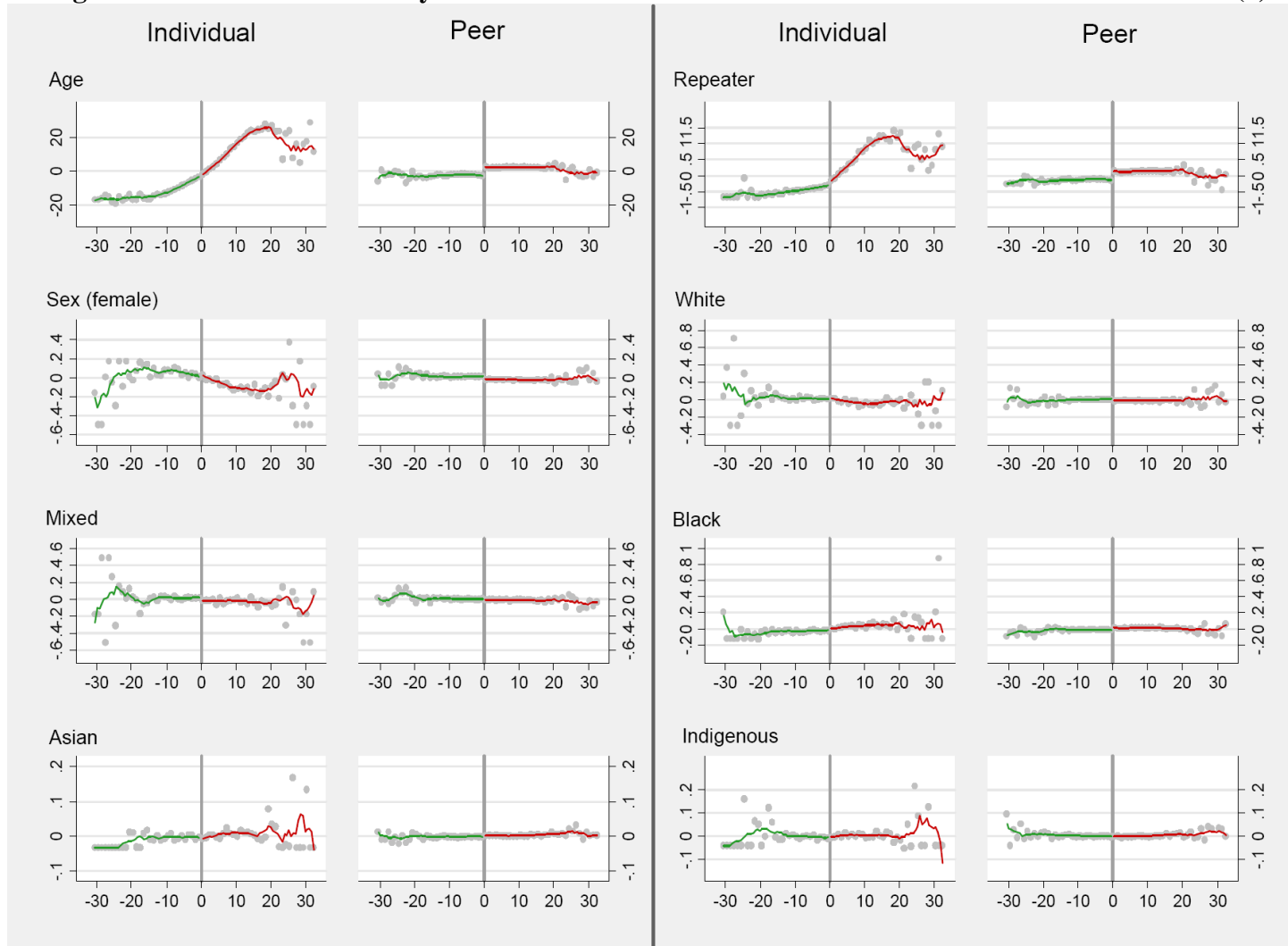
Notes: Entries are means of the standardized categorical answers to the student questionnaire aggregated on the school level. The data comes from the student questionnaire of PROEB 2007. Standard deviations reported in parenthesis.

**Table 26: RD Estimates of Predetermined Individual Characteristics of the 2007 Entry Cohort**

Sex	-0.069	(0.070)
White	0.006	(0.087)
Mixed	-0.047	(0.053)
Black	0.050	(0.060)
Asian	0.024	(0.034)
Indigenous	0.004	(0.006)

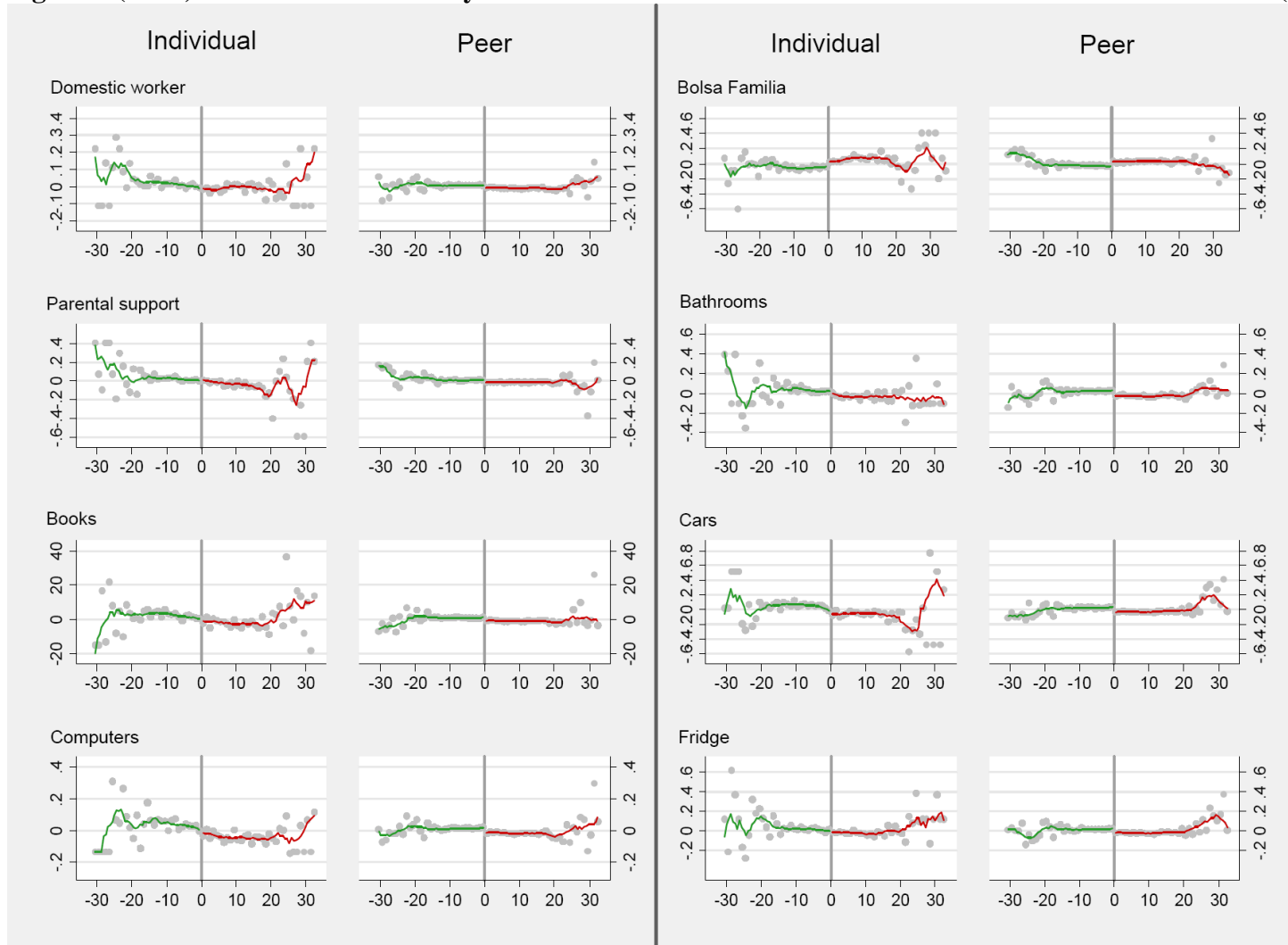
Notes: Entries are separate IV estimates of the class effect on student characteristics of first grade students of the school entry cohort of 2007, where being in the second class has been instrumented by a dummy for having an age rank larger than 0. The data comes from official records of the 2007 school census. For each variable a separate regression has been estimated. All specifications include a second-order polynomial in the age rank of students. Heteroskedasticity consistent standard errors, clustered on the school level are reported in parentheses.

**Figure 8: Test for Discontinuity of Individual and Peer Values of Pre-determined Characteristics (1)**



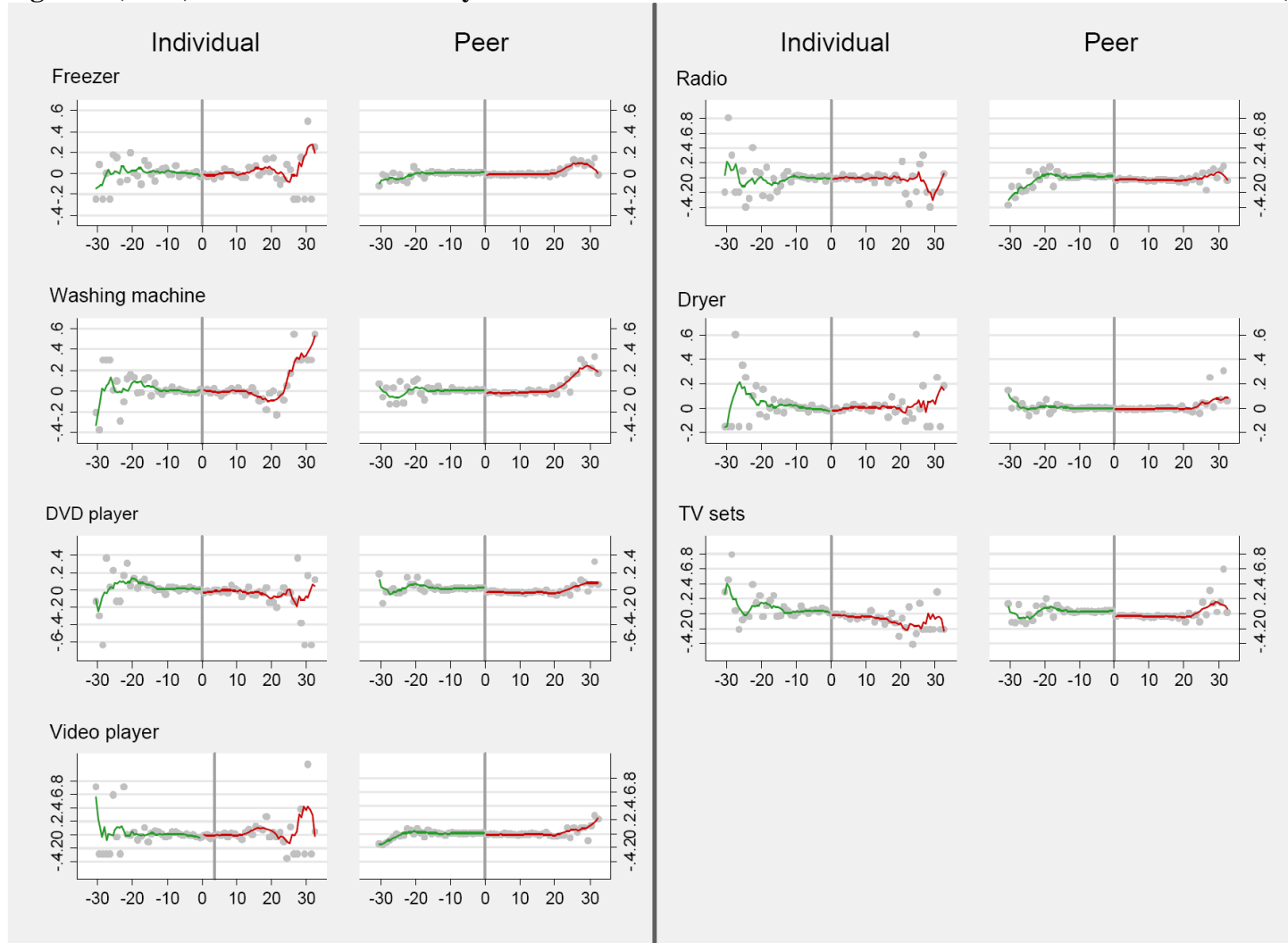
Notes: The graphs plot local averages of individual values (columns 1 & 3) and of the value for the peers of the individual students (columns 2 & 4) according to the age ranking in the cohort as distance of students from the cut-off point and local linear regression fits on both sides of the cut-off point using a rectangular kernel with a bandwidth of 3 months.

**Figure 8 (cont.): Test for Discontinuity of Individual and Peer Values of Pre-determined Characteristics (2)**



Notes: The graphs plot local averages of individual values (columns 1 & 3) and of the value for the peers of the individual students (columns 2 & 4) according to the age ranking in the cohort as distance of students from the cut-off point and local linear regression fits on both sides of the cut-off point using a rectangular kernel with a bandwidth of 3 months.

**Figure 8 (cont.): Test for Discontinuity of Individual and Peer Values of Pre-determined Characteristics (3)**



Notes: The graphs plot local averages of individual values (columns 1 & 3) and of the value for the peers of the individual students (columns 2 & 4) according to the age ranking in the cohort as distance of students from the cut-off point and local linear regression fits on both sides of the cut-off point using a rectangular kernel with a bandwidth of 3 months.

### **3. The Effect of Violence on Birth Outcomes**

#### **3.1 Introduction**

In this project we analyze birth outcomes of children whose mothers were exposed to high levels of violence in their local environment during pregnancy. There is considerable evidence showing that the nine months in utero are critical in shaping a person's life, affecting a variety of economic and non-economic outcomes even in adulthood. Although there is a small but growing literature in economics showing that maternal stress and exposure to extreme events, including conflict and terrorist attacks, during pregnancy affect birth outcomes, the impact of day-to-day violence is, by and large, understudied.

Exposure to violence in utero might affect birth outcomes directly through the mother's fear of victimization and psychological stress, which is in turn known to lead to worse birth outcomes. Violence can also affect mothers and hence the health of the fetus directly through victimization, with its ensuing negative economic, physical, and psychological consequences. Indirect effects, such as changes in labor supply, might also be at play, with effects on household income, increased difficulties in, or higher costs of, accessing local health services due to safety concerns, or even changes in fertility, possibly affecting observed birth outcomes through selection. Additionally, resource diversion on the part of both households and communities in order to prevent or counteract violence might lead to reductions in expenditures associated with children's well-being.

This analysis focuses on Brazil, a country with one of the highest levels of violence worldwide (UNODC, 2011), with a homicide rate of 21 deaths per 100,000 population (as of 2011), approximately five times the rate in the United States and almost 20 times the rate in the United Kingdom. Homicide is the leading cause of death in men aged 15-44 (Reichenheim et al., 2011), and day-to-day violence is a top concern

among citizens of Brazil. In the 2010 Latinobarometer, about 16 percent of Brazilian respondents listed violence and public security as the most important problem (Latinobarometer, 2010), and existing estimates put the direct costs of violence and crime at between 3 and 5 percent of annual GDP (Couttolene, Cano, Piquet Carneiro, and Phebo, 2000; Kahn 1999; Heinemann and Verner, 2006; Velasco Rondon and Viegas, 2003; World Bank, 2006).<sup>60</sup>

In order to assess the impact of violence on birth outcomes, we combine microdata on all births for 11 years (2000 to 2010) from official birth records with information on all homicides that occurred over the same period obtained from official death records. Vital statistics provide the date of birth and the place of residence of the mother up to the municipality level. Similarly, for homicides, the data provide information on the date and municipality of occurrence of the death. This allows us to identify the incidence of homicides during different stages of pregnancy in the mother's municipality of residence.

Homicide rates are often used as crime and violence indicators (UNODC, 2011).<sup>61</sup> Evidence for Brazil, in particular, shows a close correlation between different forms of violent crime and homicides (World Bank, 2006).<sup>62</sup> Because of their severity, underreporting is not generally a concern (Heinemann and Verner, 2006), and homicides are more likely to be followed up by police investigations and media coverage relative to other types of crime, making them particularly visible to the public. As uniform crime reports are not publicly available for Brazil, homicide rates from

---

<sup>60</sup> Methodologies such as contingent valuation surveys and willingness-to-pay methods (see Soares 2010 for a description of the methods and a survey of the findings) have not been applied in the Brazilian context.

<sup>61</sup> In order to measure local violence we use local homicide rates. Homicides data are considered among the most representative and comparable crime and violence indicators and serve as a frequently used and reasonable proxy for violence (UNODC 2011). In contrast, other forms of violence that may be reported in surveys are subject to biases as a measure of the severity of violent episodes may depend on subjective perceptions.

<sup>62</sup> A significant proportion of murders in Brazil is associated with drug trafficking and the ensuing disputes over territory, distribution, and leadership (UNODC, 2005). Murders based on drug trafficking—but not exclusively those—are related to a wide variety of other violent activities, such as robberies, kidnapping, assaults, and muggings (Heinemann and Verner, 2006).

death records constitute a unique source of information on violence that is uniform across space and time.

The rich information available in the vital statistics data allows us to measure the effects of violence on a variety of outcomes, including birthweight, APGAR scores, gestational length and infant mortality, as well as potential margins of selection due to fertility, abortion, and miscarriage.

Identification is based on a difference-in-differences strategy across geographical areas and time (conditional, in some specifications, on municipality linear trends). This allows us to obtain credible estimates of causal impact and provides the opportunity for a falsification test. The sheer amount of data helps us obtain precise estimates: this is crucial, as some of these phenomena (e.g., infant mortality) are rare events and their statistical—although not necessarily their economic—magnitude may be very small and hard to detect in sample surveys. Using information on gestation allows us to reconstruct the date of conception rather than relying on date of birth. This allows us to obtain estimates of the impact of homicides on birth outcomes (e.g. birthweight) that are correlated with length of gestation and that are free of potential selection bias when using date of birth.

Most of the analysis focuses on small – primarily rural – municipalities (with populations of less than 5,000), for which municipality-level homicide rates provide a localized measure of violence. These small municipalities form a more homogeneous group of municipalities and for these it is more credible that the homicide rate is a measure of local violence and that it exerts enough variation over time.

To preempt our results, we show that in small municipalities, one extra homicide during pregnancy leads to an increase in the probability of low birthweight (<2.5 kg.) of around half a percentage point, a 6 percent increase relative to baseline (8 percent). Consistent with findings elsewhere in the literature, the effect seems to be concentrated



in the first trimester of pregnancy. The estimated effect is economically meaningful, being approximately ten times the effect estimated for the United States of being a recipient of Food Stamps (Almond, Hoynes, and Whitmore Schanzenbach, 2011) (although clearly a much larger fraction of households are in receipt of Food Stamps compared to those exposed to homicides). The effect seems to be largely driven by increased prematurity rather than intrauterine growth retardation. We find no effect on child mortality or margins of endogenous fertility.

### **3.2 Birth Outcomes and In Utero Experiences: The Effect of Exposure to Violence**

The consequences of low birthweight and fetal health more generally on long-run outcomes, such as educational attainment, later life health, mortality, and labor market performance have been established in a large body of literature (Alderman and Behrman, 2006; Almond, Chay, and Lee, 2005; Almond and Currie, 2011b; Currie, 2011; Currie and Moretti, 2007; Royer, 2009; Victora, Kirkwood, Ashworth, Black, Rogers, Sazawal, Campbell, and Gore; 1999). Low-birthweight infants display a substantially increased risk of neonatal or infant death and are more likely to require additional outpatient care and hospitalization during childhood, adding to the private and social costs of poor birth outcomes. Of those living into adulthood, some may suffer from cognitive and neurological impairment, conditions typically associated with lower productivity in a range of educational, economic, and other activities, as well as from increased morbidity (e.g., risk of cardiovascular disease, diabetes, and hypertension).

The importance of fetal shocks and of the circumstances in utero on birth and later life outcomes has only been recently acknowledged by economists, leading to resurgent interest both in the theoretical and in the empirical literature. There are now numerous empirical studies showing that, consistent with Barker's fetal origin hypothesis, the nine months in utero constitute a critical period of a person's life,

shaping subsequent health, educational, and labor market outcomes (Almond and Currie, 2011a; Almond and Currie, 2011b; Currie, 2011).

Almond and Currie (2011a) categorize factors affecting the prenatal environment into three groups: those affecting maternal and thereby fetal health (such as nutrition and infection), economic shocks, and environmental pollution. A number of studies, in particular, have established a link between household maternal nutrition and birth outcomes, especially birthweight, one of the most important and easiest to measure predictors of economic and non-economic outcomes in adulthood. Some studies focus on the role of redistributive policies (see, for example, Almond, Hoynes, and Whitmore Schanzenbach, 2011 on the U.S. Food Stamps program and Amarante, Manacorda, Miguel, and Vigorito, 2011 on the Uruguayan PANES), while others focus on the role of famines, natural disasters, or even fasting during pregnancy (Almond, 2006; Almond and Mazumder, 2011; Banerjee, Duflo, Postel-Vinay, and Watts, 2010). For Brazil, Rocha and Soares (2012) show that negative weather shocks during pregnancies lead to a significant reduction in gestational length and birthweight. Other studies focus instead on the disease environment during pregnancy (see Almond, 2006 and Kelly, 2011 on maternal influenza and Barreca, 2010 for maternal exposure to malaria) and on pollution (Currie and Walker, 2011; Chay and Greenstone, 2003 on air pollution, Almond, Edlund and Palme, 2009 on nuclear fallout, and Reyes, 2007 and Nilsson, 2011 on leaded gasoline), showing that both play substantial roles in affecting birth and later outcomes.

Despite evidence that maternal stress during pregnancy negatively affects cognition, health, and educational attainment of children through elevated levels of the stress hormone cortisol (Aizer, Stroud, and Buka; 2009), presumably because of data limitations, the effect of exposure to crime and violence on birth outcomes has received considerably less attention.

A stream of literature focuses on terrorist attacks and conflict. Camacho (2008) finds that landmine explosions during the first trimester of pregnancy have a significant negative effect on birthweight in Colombia, with one extra landmine explosion during pregnancy leading to a decrease in birthweight by 8.7 grams. Ecclestone (2012) shows that exposure to the 9/11 terror attacks among pregnant women in New York City led to a reduction in birthweight of between 12 and 14 grams and an elevated level of prematurity. In a setting closer to ours, Mansour and Rees (2012) find a modest but imprecisely estimated increase in the fraction of low birthweight infants in response to an increase in noncombatant fatalities in the West Bank and Gaza during the second Intifada.<sup>63</sup>

Although clearly related to this chapter, these studies focus on the effect of rare, extreme events, implying that their findings may not necessarily be applicable in other settings where violence is endemic.

### 3.3 Background, Trends, and Data

#### 3.3.1 Births and Birth Outcomes

In order to characterize the distribution of birthweight and other birth outcomes, in this chapter we use microdata from birth certificates, which are collected by the Brazilian Ministry of Health through *DATASUS*, literally the *Departamento de Informática do Sistema Único de Saúde* (SUS).<sup>64</sup> The data provide a large array of information on pregnancy and newborns' outcomes as well as on mothers' characteristics. Coverage is

---

<sup>63</sup> There is very little evidence on the effect of mother's victimization. One exception is Aizer (2011), which shows that mother's domestic-violence-induced hospitalization considerably reduces birthweight.

<sup>64</sup> The information on births is first collected by the health institution where the birth took place and then forwarded to the state's health secretariat (via means of the municipal health secretariat), which in turn is responsible for entering the information into the central database (FUNASA 2001). In the rare case of a home birth, this information is submitted by medical staff attending the birth.

practically universal: data from the 2010 population census show that more than 99 percent of children born between 2000 and 2010 indeed have birth certificates.

Summary statistics for the period 2000-2010 are reported in the top panel of Table 27. The data provide information on more than 30 million births over the period. As said, the primary units of observation in the analysis are municipalities, relatively small geographical units roughly equivalent to a U.S. county. In the table we have information on 5,508 municipalities.<sup>65</sup> At total population of just over 181 million, each of these municipalities accounts on average for 33,000 individuals. Obviously, however, population size varies tremendously across municipalities: while São Paulo and Rio de Janeiro account for more than 11 and 6 million inhabitants respectively, more than 20 percent of municipalities have fewer than 5,000 inhabitants. For this reason, in the table we present results for all of Brazil (column (1)) and separately for the different classes of municipalities based on population size. For this we use the standard classification from the National Statistical Office (IBGE) (population 1 to 5,000; 5,001 to 20,000; 20,001 to 100,000; 100,001 to 500,000 and 500,001 or more). Smaller municipalities account for around 2 percent percent of all births.

The table illustrates that, with an incidence of low birthweight (less than 2.5 kg.) of around 8 percent, Brazil ranges above the average for OECD countries but considerably below the highest rates in some low-income countries (UNICEF, 2006). Around respectively one and half a percent of children are born with very low (<1.5 kg.) and extremely low (<1 kg.) birthweight. The data also provide information on APGAR scores, gestational length, gender, race, and a number of mother characteristics. Roughly speaking, birth outcomes are worse the greater the municipality size, although children in very large municipalities (>500,000) seem to perform better than children in large municipalities (100,000 to 500,000) among a number of dimensions.

---

<sup>65</sup> We have excluded the few municipalities that split into newer municipalities between 2000 and 2010.

Figure 9, left-hand side panel, reports average (across the entire period) low-birthweight rates in all Brazilian municipalities: darker areas correspond to municipalities with greater incidence of low birthweight. The municipalities with the highest rates of low birthweight are clustered mainly in a number of states, Maranhão and Amapá, in the Northeast and North, respectively, as well in the Southeastern states of Minas Gerais and São Paulo and Rio Grande do Sul in the South.

### 3.3.2 Infant Mortality

The middle panel of table 27 reports data on infant mortality. Data come from death certificates, which are also collected by *DATASUS*, and record very detailed causes of death, including non-natural deaths classified as homicides that we use below, as well as the date and municipality of occurrence of the death. The data also provide information on infant mortality. Infant mortality data refer to children born alive for which a birth certificate has been produced, and hence exclude fetal deaths.

The data allow us to estimate four rates: early neonatal mortality (within seven days since birth), neonatal mortality (within 28 days since birth), perinatal mortality (within the first 22 weeks since birth) and infant mortality (within the first year since birth). At nine deaths per 1,000 children, early neonatal mortality accounts for the bulk of deaths within the first year of life. Infant mortality is on average 14 per 1,000 children. Again, there is a clear gradient across municipalities, with larger municipality size being associated with worse outcomes, and with very large municipalities being somewhat below trend.

### 3.3.3 Homicides

The third panel of Table 27 reports data on homicides. These and all other aggregate statistics in the rest of the table that vary only by municipality and time are weighted by the number of births, meaning that municipality  $X$  times mean receives a weight proportional to the number of births in that cell.

For the period 2000-2010, more than 528,000 homicides are recorded, equivalent to a *yearly* homicide rate of around 26 per 100,000 individuals. Again, homicide rates tend to be higher the larger the municipality. Still, even in small municipalities, the homicide rate is 9 per 100,000 individuals. The data also provide location of death. This can be in a health institution, in one's home, in the street, or elsewhere. Clearly, when the death occurred in a health institution, the homicide might have been committed elsewhere, possibly even in another municipality, inducing considerable error in the measure of local violence that we use. The subsequent rows of the table show that around 40 percent of deaths resulting from homicides happen in the street and around 50 percent either in the street or in one's residence. Interestingly, the latter is only 44 percent in very large municipalities, where hospitals are typically located. This suggests that a fraction of homicides for which the death occurs in hospitals are likely to be committed in other municipalities. Because of this, in most of the analysis we focus on homicides for which the death occurred *in the street*. These are also likely to be the most visible and hence stress-inducing homicides, which might possibly affect pregnancy outcomes.

The middle panel of Figure 9 reports the distribution of homicide rates (in the street) across Brazilian municipalities. Municipalities with higher incidence of homicides are in the more densely populated and more urbanized areas along the coastline, as well as in the state of Bahia in the Northeast of Brazil. Municipalities with

high rates of homicides are also to be found in the less densely populated states of Mato Grosso and Pará, covering some of the Amazon region.

Although this is not immediately evident in Figure 9, once differences in population size across municipalities are taken into account, a clear positive correlation between local homicide rates and low birthweight emerges. This is shown in Figure 10, left hand-side panel, which plots the cross-sectional relationship between the fraction of low-weight births and the annual homicide rate (in the street) across all Brazilian municipalities. A predicted regression line is also superimposed and larger circles correspond to larger cities. The data clearly show that, across municipalities, higher homicide rates are associated with worse birth outcomes: the estimated coefficient is 1.6 per thousand births and highly significant at conventional levels, implying that one extra homicide out of 100,000 people is associated with 1.6 extra low-weight births out of 1,000 births. It is also clear that larger municipalities tend to outperform smaller municipalities along both of these dimensions. One possible interpretation of these correlations is that higher homicides rates are responsible for worse birth outcomes. This conclusion may be unwarranted, as different municipalities vary in characteristics which are potentially associated with both birth outcomes and mortality rates.

Indeed, the bottom part of Table 27 shows that municipalities of different sizes vary along a number of dimensions, such as income, literacy rate, and rates of urbanization. These data (like most of the municipality-level data that we use in the regressions) come from decennial population censuses.<sup>66</sup> There is evidence that larger municipalities outperform smaller ones in many socioeconomic dimensions, such as literacy rate and per capita income.

---

<sup>66</sup> The majority of the data come from population census micro-data. Additional variables have been obtained from *DATASUS*.

(<http://www2.datasus.gov.br/DATASUS/index.php?area=0206andVObj=http://tabnet.datasus.gov.br/cgi/deftohtm.exe?ibge/censo/cnv/crianpobr>). Data are available for 2000 and 2010 and we have then interpolated linearly across these two dates to estimate their value in every intervening month.

Differences in socioeconomic status and living standards across areas are also evident in the right-hand side panel of Figure 9, which displays average household income by municipality, with higher income being represented by darker areas. The Southeastern states of São Paulo, Rio de Janeiro, and parts of Minas Gerais are those with the highest average household income.

### **3.4 Econometric Methodology**

As already emphasized, the difficulty in estimating the causal effect of violence on birth outcomes is that characteristics of different residential areas are unobservable to the econometrician. Some of these unobservable characteristics might be correlated with both newborns' health outcomes and homicide rates, even in the absence of a causal effect of violence on birth outcomes. For example, children born in poorer areas are more likely to display negative birth outcomes due to the lower socioeconomic characteristics of their parents or worse provision of health services in their neighborhood, and, possibly, to be exposed to a higher (or lower) degree of violence. In this case, one would erroneously conclude that higher homicide rates lead to worse (or better) birth outcomes, a classic case of failed inference based on observational data.

In order to circumvent this problem, we propose to use a simple difference-in-differences identification strategy that relies on differential changes in homicide rates across municipality and time: this provides a way to control for unobserved time-invariant municipality characteristics and to subsume aggregate time effects. The identification strategy exploits the change in homicide rates over time while controlling for municipality and time fixed effects; because there is imperfect control over the timing of conception and no foresight over future homicides the change in the homicide rate can be considered conditionally random.

In formulas we estimate the following model:



$$Y_{mt} = \beta_0 + \beta_1 HOM_{mt} + X_{mt} \beta_3 + d_m + d_t + u_{mt} \quad (6)$$

where  $Y_{mt}$  is the average outcome variable (birthweight, still birth, infant mortality, APGAR scores, gestational length, etc.) in municipality  $m$  at time  $t$ ,  $HOM_{mt}$  is the local homicide rate,  $X_{mt}$  are vectors of average (across all individuals in each cell) individual characteristics as well as time-varying municipality-level characteristics,  $d_m$  and  $d_t$  are respectively municipality and time-fixed effects, and  $u$  is an error term. We estimate equation (6) on aggregate month X municipality level data, which is the level of variation of the homicide data (rather than on individual data), for computational purposes. All regressions are estimated using WLS, with weights given by the number of births in each cell.

In the empirical analysis, we estimate the effect of the homicide rate at different stages of pregnancy (i.e., first, second, and third trimester) and test for the validity of the identification assumption by introducing in the regressions additional pre- and post-pregnancy homicide rates as additional regressors. One would expect homicide rates pre- and post-pregnancy not to affect birth outcomes: finding a significant coefficient on the latter would point to a violation of the identification assumption.

In the following, we measure trimesters of pregnancy starting from the date of conception. We recover the latter based on the child's date of birth minus the length of gestation. As the length of gestation is recorded in intervals in our data (<22, 23-27, 28-31, 32-36, 37-41, >41 weeks), we use the mid-point of each interval. This approach has multiple advantages. First, it allows us to correctly measure exposure in different trimesters of pregnancy, which would not be possible if we counted retrospectively since the time of birth (as typically done in this literature) and ignored the variation in the length of gestation across pregnancies. Second, it allows us to directly estimate the effect of homicides on the length of gestation, a potentially interesting outcome in itself. Third, and related to the latter, it allows us to obtain estimates of the impact of

homicides on other outcomes (e.g. birthweight) that are correlated with length of gestation and that are free of potential selection bias.

### 3.5 Empirical Results

#### 3.5.1 Birthweight

Table 28 presents estimates of equation (6) for small municipalities (<5,000 individuals). Small municipalities are concentrated in a few states (Tocantins, Piauí, Goiás, Minas Gerais, São Paulo, and Rio Grande do Sul) and geographically rather dispersed (see Figure 12). The table reports results on average birthweight (in grams) and on the fraction of low, very low, and extremely low-weight births (per 1,000 births). Column (1) of Table 28 reports a simple difference-in-differences estimate for the effect of the homicide rate on average birthweight in the first, second, and third trimesters since conception. Regressions include only municipality and month of conception fixed effects. Homicide rates here are computed at the quarterly level (i.e., number of homicides per quarter over total population). The data show a negative and very precisely estimated effect of the homicide rate in the first trimester of gestation on birthweight. The estimated effect of an increase by one in the number of quarterly homicides per 100,000 individuals in column (1) is just below half a gram (-0.43 grams.) This implies that in an average municipality in this class (around 3,700 individuals), one extra homicide will lead to a reduction in average birthweight among children exposed to that homicide in their first trimester of pregnancy of around 12 grams ( $= (0.43 \times 100,000) / 3,700$ ). This is a small effect, on the order of 0.4 percent relative to an average birthweight of 3.210 kg. For comparison, for Colombia, Camacho (2008) finds that one landmine explosion during early pregnancy reduces birthweight by 7.5 grams.

The estimates for the second and third trimester are positive, much smaller in magnitude but not significant at any conventional levels. This is in line with findings elsewhere in the literature that stress induced by extreme events matters mostly during the first trimester of gestation (Camacho, 2008; Torche, 2011; Mansour and Rees, 2012).

Column (2) controls for a very rich set of mother and child characteristics and time-varying municipality characteristics from census data (see notes to Table 28), including municipality-specific linear time trends and municipality x calendar month (January to December) effects. Results are essentially unchanged relative to column (1), lending credibility to the identification assumption that—conditional on time- and municipality-fixed effects—the variation in the homicide rate across municipalities and time is almost as good as random. In column (3), we additionally include homicide rates in the fourth, fifth, and sixth trimester since conception, that is - for pregnancies of normal gestational length - in the first, second, and third trimester since birth and homicide rates in the three trimesters before conception. The inclusion of these variables makes virtually no difference to the results while we find no significant coefficients on the different lead and lag variables, lending support to our identification assumption. Figure 11 plots the point estimate and 95 percent confidence interval of the effect of homicide rate on birthweight and low birthweight for the three trimesters prior to conception, and the six trimesters after conception. Only the coefficient for the first trimester in utero reveals a significant negative effect, while the point estimates for the other quarters are much smaller and not significant.

Results in the following columns of the Table confirm these findings and show that homicides have a particularly pronounced effect at the bottom tail of the birthweight distribution. We find significant effect of homicides in the first trimester the fraction of low, very low, and extremely low-weight births of, respectively, 0.17, 0.06,

and 0.04 per 1,000 births. In turn this means that one extra homicide in a small municipality will lead to an increase in the fraction of low, very low, and extremely low birthweight children of 0.5 ( $= (0.17/10) \times 100,000 / 3,700$ ), 0.2 and 0.1 percentage points, that is, respectively a 6, 16, and 21 percent increase (relative to a baseline incidence of 0.078, 0.010, and 0.005).

### 3.5.2 Additional Outcomes

Table 29 reports regression results on a number of additional outcomes. For brevity, we report only specifications with the entire set of controls as in column (2) of Table 28. Column (1) reports the effect of homicides on gestational length. Indeed, homicides in the first trimester increase prematurity, by lowering gestational length. Column (2) reports the effect on APGAR scores. We use the average score one minute and five minutes post birth in an attempt to boost precision: still we find no evidence of a significant effect of increased levels of violence on this outcome.

Columns (3) to (6) report the effects on mortality rates at different intervals since birth. The dependent variable here is the fraction of deaths per thousand children born alive. Again, there is no evidence of violence affecting child mortality rates.

Columns (7) and (8) report estimates of impact on birthweight and low birthweight only for pregnancies of normal gestational length, defined as pregnancies of 37 weeks or more. We report results on birthweight and the fraction of low-birthweight children (as in columns (2) and (5) of Table 28). Interestingly, results on birthweight disappear. Combined with the findings in column (3) of Table 29, this suggests that violence leads to greater rates of prematurity and, via this, to increased risk of low birthweight.

The last concern we have pertains to selective fertility. Violence might affect birth outcomes through the selection that it operates on the number of children who are

eventually born. This can happen through a variety of margins: selective sexual activity or contraception use, selective fetal mortality, abortion, and miscarriage. In order to study these combined effects, in column (9) of Table 29 we report a regression of the log number of births by municipality and time on the same variables as in columns (1) to (8) with the exception of mother characteristics. As in the other regressions, we control for the age and gender structure of the population in each municipality X time cell. The latter allows us to control for differences in the population at risk (women of fertile age) across cells. We find very small and statistically insignificant effects on fertility, implying that selection along this margin is unlikely to explain our results.

### **3.5.3 Alternative Definitions of Homicide**

In Tables 28 and 29, we restrict reporting to homicides for which the death occurred in the street. Table 30 reports results using, respectively, homicides in the street and in one's residence (columns (1) to (5)) and all homicides, that is, also those for which death occurred in health institutions (columns (6) to (10)). Using additionally homicides for which the death occurred in residences makes virtually no difference to our results. Estimates, however, become smaller and less precise when we use all homicides: this is consistent with the notion that homicides for which the death occurred in hospital provide an error-ridden measure of local violence.

### **3.5.4 Heterogeneous Effects by Mother's Education**

To conclude, in Table 31 we report separate regression results for infants born to mothers with incomplete and complete primary education (8 years of schooling) respectively. Each of these two groups account, roughly, for 50 percent of births. The effect seems to manifest largely among children of poorly educated mothers. Although

results for highly educated mothers are qualitatively similar, point estimates are typically smaller and statistical significance is lower. It appears that violence adds up to the disadvantage that children of poorly educated mothers already suffer as a result of their household's lower socioeconomic status.

### **3.6. Summary of Findings and Conclusions**

Using a very rich dataset on the universe of births and homicides from vital statistics data over the period 2000-2010, we estimate the effect of in-utero exposure to homicides on a range of birth outcomes in small Brazilian municipalities. We find a significant negative effect of exposure to violence during the first trimester on birthweight, which is in line with findings on the effect of other stress-related shocks during pregnancy in the literature. We also find significant and large positive effects of homicides on the probability of low birthweight, implying that the effects are particularly pronounced at the bottom tail of the birthweight distribution. Our results are robust to the introduction of maternal and municipal socioeconomic controls, including municipality-specific linear time trends. A falsification exercise, consisting of testing for the effect of pre- and post-pregnancy homicide rates on birth outcomes, lends further credibility to our identification assumption.

We show that violence in the first trimester of pregnancy affects birth outcomes through reduced gestational length. Increased prematurity hence, rather than intrauterine growth retardation, seems to explain the pronounced effect on low birthweight that we have documented in the chapter.

As violence might affect the probability of appearing in the data set through changes in fertility or possibly via abortion or miscarriage, one concern is that our results might be driven by selection. That is, there may be a differential response to increased levels of violence among women with differential propensity to give birth to

low-weight infants. Despite this concern, we find no evidence of homicides affecting fertility outcomes.

Finally, we show that results are largely concentrated among poorly educated mothers, that is, those with less than completed primary education. This suggests that violence adds up to the mechanisms that affect the transmission of socioeconomic status between parents and their offspring.

Although our estimates for the effect of one extra homicide in small municipalities are economically meaningful, high homicide rates are not responsible for the high level of low birthweight in Brazil. This is because overall, homicides are rather rare events. At current rates, and if one is willing to extrapolate the estimates from small municipalities to the whole of Brazil, our back-of-the-envelope calculations show that homicide rates account only for a minimal fraction (0.01%) of total low-birthweight incidence in the country.

**Table 27: Descriptive Statistics**

	All	By municipality size				
		<5000	5,000-19,999	20,000-99,999	100,001-500,000	>500,000
Number of municipalities	5,508	1,289	2,648	1,320	215	36
Number of births	30,367,939	616,733	4,491,073	8,808,710	7,254,770	9,106,653
Birthweight	3184.190	3,210.62	3,222.315	3,207.675	3,164.246	3,156.751
Low birthweight	0.087	0.078	0.080	0.082	0.092	0.082
Very low birthweight	0.012	0.010	0.011	0.011	0.013	0.011
Extremely low birthweight	0.006	0.005	0.005	0.005	0.006	0.004
Gestational length	38.690	38.751	38.755	38.748	38.659	38.622
APGAR – 1 minute	8.144	8.143	8.073	8.109	8.185	8.176
APGAR – 5 minutes	9.235	9.300	9.226	9.242	9.250	9.216
Female	0.512	0.514	0.513	0.513	0.512	0.512
White	0.502	0.598	0.471	0.459	0.561	0.508
Prenatal visits	5.705	5.803	5.446	5.458	5.889	5.920
Mother's age	26.168	26.022	26.223	25.744	25.933	26.754
Mother never married	0.613	0.563	0.633	0.657	0.622	0.601
Mother's years of schooling	7.826	7.745	7.256	7.695	8.865	7.736
Early neonatal mortality (1 wk.)	9.121	8.000	8.832	9.110	11.881	7.042
Neonatal mortality (4 wks.)	11.211	9.767	10.625	10.901	14.486	9.184



Table 27 continued

Perinatal mortality (22 wks.)	13.595	11.973	13.129	13.441	16.923	11.316
Infant mortality (1 year)	14.706	12.951	14.446	14.746	17.986	12.170
Homicide rate	26.284	9.102	12.832	19.381	32.650	36.613
Homicide rate, in the street	10.972	2.498	4.284	7.903	14.574	15.330
Homicide rate, in the street and in homes	13.888	4.691	6.855	10.794	17.999	18.067
Population	1,170,281	3,703	12,638	49,120	250,081	3,887,465
Urbanization rate	0.822	0.531	0.574	0.721	0.937	0.984
HH income 2010 \$R	1,100.41	582.53	571.05	752.81	1,150.36	1,663.70
Literacy rate	0.817	0.758	0.723	0.755	0.848	0.878

Source: *DATASUS* and IBGE population census.

Notes: All entries are weighted by the number of births. Neonatal and infant mortality rates are expressed as a fraction per 1,000 live births. Homicide rates are expressed as a fraction per 100,000 population.

**Table 28: The Effect of Homicides during Pregnancy on Birthweight by Trimester since Conception – Small Municipalities**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Trimester	Birthweight (grams)			Low birthweight (x 1,000)			Very low birthweight (x 1,000)			Extremely low birthweight (x 1,000)		
1 (pre-conception)			-0.0439 (0.1037)			0.0011 (0.0480)			-0.0161 (0.0184)			-0.0081 (0.0121)
2 (pre-conception)			-0.1344 (0.1090)			0.0621 (0.0531)			0.0322 (0.0232)			0.0279* (0.0151)
3 (pre-conception)			-0.0954 (0.1021)			0.0036 (0.0494)			0.0006 (0.0191)			-0.0123 (0.0111)
1	-0.4328** (0.1722)	-0.4506*** (0.1627)	-0.4551*** (0.1625)	0.1488* (0.0890)	0.1680** (0.0836)	0.1709** (0.0839)	0.05563 (0.0354)	0.0580* (0.0344)	0.0593* (0.0342)	0.0403 (0.0258)	0.0393 (0.0256)	0.0416* (0.0252)
2	0.0192 (0.2030)	0.0568 (0.1918)	0.0507 (0.1920)	0.0688 (0.0956)	0.0584 (0.0890)	0.0624 (0.0891)	-0.0310 (0.0342)	-0.0356 (0.0337)	-0.0336 (0.0337)	-0.0149 (0.0240)	-0.0209 (0.0242)	-0.0183 (0.0240)
3	0.0214 (0.1967)	-0.0486 (0.1885)	-0.0523 (0.1904)	-0.0272 (0.0882)	0.0038 (0.0859)	0.0059 (0.0862)	0.0399 (0.0415)	0.0407 (0.0408)	0.0418 (0.0414)	0.0221 (0.0315)	0.0194 (0.0305)	0.0212 (0.0313)
4 (post-birth)			-0.0946 (0.1802)			0.0513 (0.0850)			0.0380 (0.0380)			0.0285 (0.0314)
5 (post-birth)			-0.0313 (0.1898)			0.0095 (0.0842)			0.0056 (0.0459)			0.0271 (0.0373)
6 (post-birth)			-0.1005 (0.2086)			0.0806 (0.0944)			0.0286 (0.0402)			0.0480 (0.0373)
Municipality f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pregnancy controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Mother controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: Each column reports the results from a separate regression of the dependent variable on the local quarterly homicide rate in different trimesters since the month of conception. Homicide rates are expressed as fraction per 100,000 individuals. Fraction birthweight is expressed per 1,000 live births. Regressions are run on cells defined by municipality and time of conception with weights equal to the number of births by cell. Controls include number of newborns by gender and race (black, white, mixed, Asian, indigenous) and number of multiple

births (twins, triplets, more than three children). Mother controls include age (10-19, 20-39, etc.), marital status (single, married, divorced, widowed), years of completed education (no education, 1-3, 4-7, 8-11, 12 and more), average number of previously born alive children and of stillbirths. Municipality controls include fraction of households with possession of radio, TV, washing machine, telephone, computer, and fraction with access to piped water, waste collection, electricity, fraction of the population by gender and age, fraction of adult population literate, average years of schooling in the population, fraction of families with *Bolsa Família*, health establishments and nurses per capita, unemployment rate, urbanization rate, fraction of children in work, interaction of municipality with calendar month and municipality trends. Clustered standard errors by municipality in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Number of cell observations: 136,711 (616,733 births).

**Table 29: Homicide Rates and Additional Birth Outcomes by Trimester since Conception – Small Municipalities**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Gestational length (weeks)	APGAR score (avg. 1 & 5 minutes)	Child mortality rates (x 1,000)				Only pregnancies of normal gestation length (37 weeks or more)		Log number of births
Trimester			Early neonatal (1 week)	Neonatal (4 weeks)	Perinatal (22 weeks)	Infant (1 year)	Weight	Low birthweight (x 1,000)	
1	-0.0011* (0.0005)	-0.1199 (0.4685)	-0.0384 (0.1131)	0.0201 (0.1345)	0.0510 (0.1558)	0.0872 (0.1599)	-0.1877 (0.1558)	0.0884 (0.0689)	0.0001 (0.0001)
2	0.0004 (0.0005)	0.1905 (0.4673)	0.1190 (0.1289)	0.0754 (0.1379)	0.1512 (0.1468)	0.1197 (0.1558)	-0.0073 (0.1860)	0.0519 (0.0731)	-0.0000 (0.0001)
3	-0.0005 (0.0005)	-0.1657 (0.4851)	0.0410 (0.1148)	-0.0415 (0.1190)	-0.0500 (0.1376)	-0.0874 (0.1330)	0.0541 (0.1796)	-0.0560 (0.0739)	-0.0001 (0.0001)
Municipality f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pregnancy controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Mother controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Municipality controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Dependent variable in columns 3 to 6 is fraction of children dead per 1,000 live births. See also notes to Table 28.

**Table 30: The Effect of Homicides during Pregnancy on Birthweight by Trimester since Conception –Alternative Definition of Homicide Rate - Small Municipalities**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Homicides in the street and in one's home					All homicides				
	Birthweight (grams)	Low birthweight (x 1,000)	Very low birthweight (x 1,000)	Extremely low birthweight (x 1,000)	Gestational length (weeks)	Birthweight (grams)	Low birthweight (x 1,000)	Very low birthweight (x 1,000)	Extremely low birthweight (x 1,000)	Gestational length (weeks)
Trimester										
1	-0.3308*** (0.1250)	0.0816 (0.0616)	0.0588** (0.0279)	0.0480** (0.0218)	-0.0007* (0.0004)	-0.1674* (0.0904)	0.0156 (0.0464)	0.0179 (0.0193)	0.0177 (0.0143)	-0.0002 (0.0003)
2	-0.0039 (0.1331)	0.0731 (0.0619)	-0.0206 (0.0247)	-0.0113 (0.0196)	0.0004 (0.0003)	-0.0133 (0.1070)	0.0259 (0.0499)	-0.0329* (0.0183)	-0.0226 (0.0141)	0.0005* (0.0003)
3	-0.0272 (0.1389)	0.0046 (0.0635)	0.0309 (0.0277)	0.0170 (0.0203)	-0.0002 (0.0003)	0.0327 (0.1003)	-0.0380 (0.0487)	0.0054 (0.0198)	-0.0019 (0.0146)	-0.0001 (0.0003)
Municipality f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pregnancy controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mother controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: See notes to Table 28.

**Table 31: Homicide Rates and Additional Birth Outcomes by Trimester since Conception – by Mother’s Education – Small Municipalities**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Incomplete primary education				Completed primary education			
Trimester	Birthweight (grams)	Low birthweight (x 1,000)	Gestational length (weeks)	APGAR score (avg. 1 & 5 minutes)	Birthweight (grams)	Low birthweight (x 1,000)	Gestational length (weeks)	APGAR score (avg. 1 & 5 minutes)
1	-0.4738** (0.2182)	0.1630 (0.1188)	-0.0011 (0.0007)	-0.2976 (0.6173)	-0.2775 (0.2608)	0.1208 (0.1221)	-0.0008 (0.0008)	0.1086 (0.5834)
2	0.1104 (0.2427)	0.1224 (0.1148)	0.0009 (0.0006)	0.1349 (0.5814)	-0.1116 (0.2748)	-0.0546 (0.1288)	0.0001 (0.0008)	0.5817 (0.5922)
3	-0.0325 (0.2521)	-0.0941 (0.1140)	-0.0003 (0.0006)	0.3337 (0.6216)	-0.1436 (0.2557)	0.1725 (0.1208)	-0.0007 (0.0007)	-0.7016 (0.5512)
Municipality f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pregnancy controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mother controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Incomplete primary education corresponds to less than 8 years of completed education. Number of observations in columns 1 to 4 is 115,922 while in columns 5 to 8 this is 109,510. See also notes to Table 28.

**Table 32: The Effect of Homicides during Pregnancy on Birthweight by Trimester since Conception – By Municipality Size**

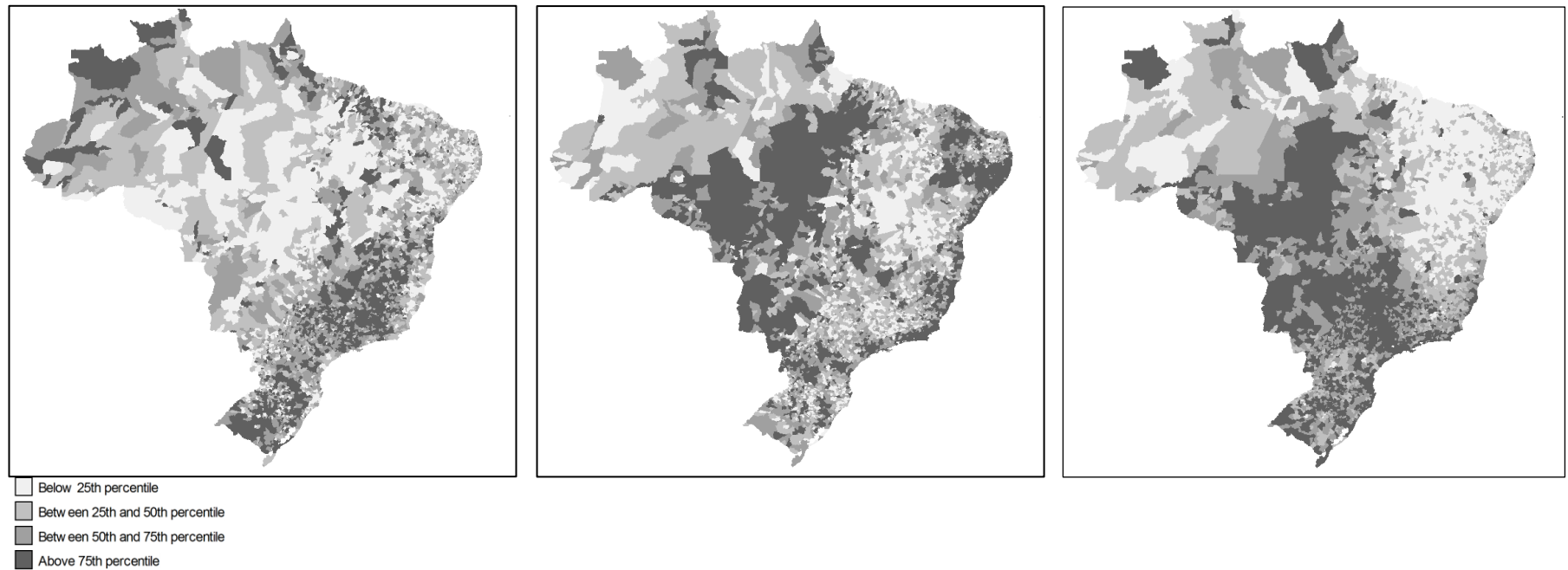
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Trimester	5,001-20,000		20,001-100,000		100,001-500,000		>500,000	
1 (pre conception)		-0.0109 (0.0648)		-0.0174 (0.0657)		0.0769 (0.1302)		-0.2723 (0.3754)
2 (pre conception)		-0.0131 (0.0573)		-0.0562 (0.0718)		0.0885 (0.1144)		-0.3816* (0.2226)
3 (pre conception)		-0.0314 (0.0605)		0.0629 (0.0792)		0.0570 (0.1208)		0.1392 (0.3214)
1	0.0911 (0.0990)	0.1296 (0.0934)	-0.1452 (0.1089)	-0.0399 (0.1038)	0.903 (0.2118)	0.0209 (0.1864)	0.3289 (0.4537)	0.9670** (0.3554)
2	-0.0702 (0.1047)	-0.0525 (0.0984)	-0.0267 (0.1112)	0.1531 (0.1042)	0.1720 (0.2118)	-0.0531 (0.1894)	0.0187 (0.3572)	0.0359 (0.3491)
3	0.0563 (0.1001)	0.0597 (0.0965)	-0.0570 (0.1155)	-0.0147 (0.0990)	-0.2779 (0.1842)	-0.2898 (0.1771)	-1.3954*** (0.4345)	-1.0908** (0.4561)
4 (post birth)		-0.0474 (0.0985)		-0.1782* (0.0948)		-0.2064 (0.1601)		0.3185 (0.4327)
5 (post birth)		0.0676 (0.0975)		-0.1233 (0.0940)		0.0590 (0.1649)		0.0417 (0.4859)
6 (post birth)		0.0686 (0.0966)		-0.0560 (0.0965)		-0.1838 (0.1473)		-0.1287 (0.5735)
Municipality f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pregnancy controls	No	Yes	No	Yes	No	Yes	No	Yes
Mother controls	No	Yes	No	Yes	No	Yes	No	Yes
Municipality controls	No	Yes	No	Yes	No	Yes	No	Yes
Number of cell observations	300,436	300,436	150,358	150,358	24,500	24,500	3,978	3,978
Number of individuals observations	4,491,073	4,491,073	8,808,710	8,808,710	7,254,770	7,254,770	9,106,653	9,106,653

Note: See notes of Table 28.

**Figure 9: Municipality Characteristics**  
 Fraction low-weight births

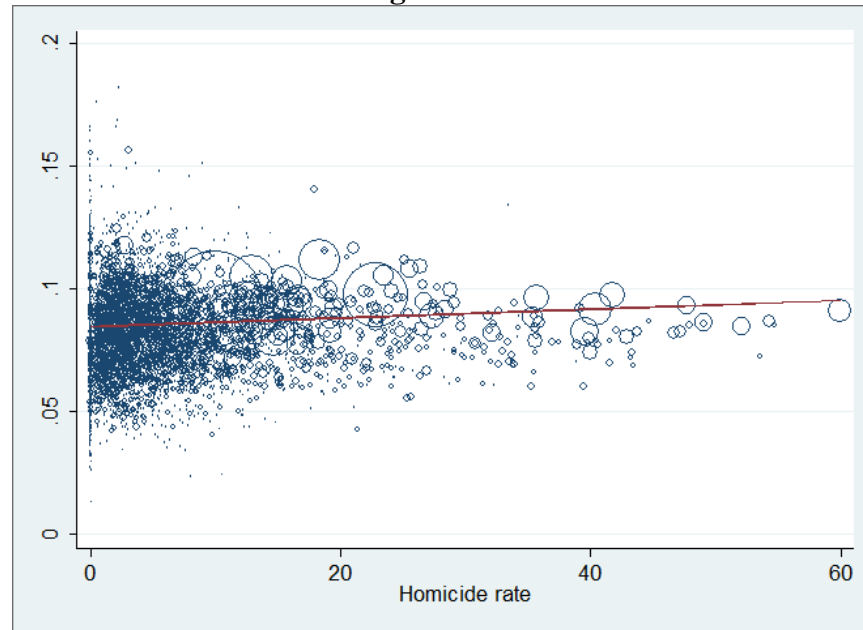
Homicide rate (in the street)

Average household income



Notes. The pictures report, respectively, the average fraction low-weight births (<2.5 kg), the homicide rate in public places, and household income between 2000 and 2010.

**Figure 10: Incidence of Low Birthweight and Homicide Rates across Municipalities**

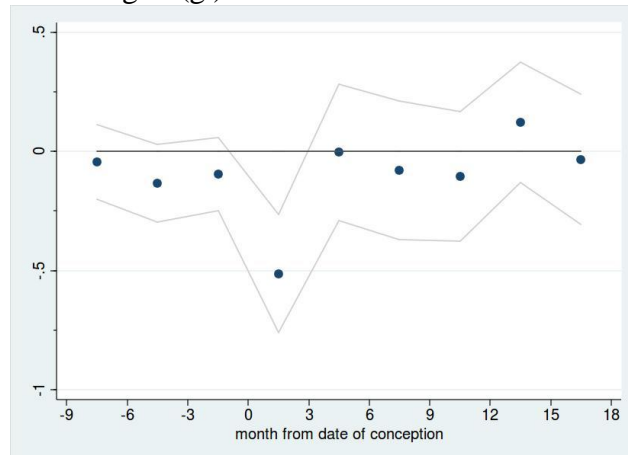


Note: The figure reports the relationship between the fraction of low-weight births and the annual homicide rate (in the street) across all Brazilian municipalities. A predicted regression line is superimposed and larger circles correspond to larger cities.

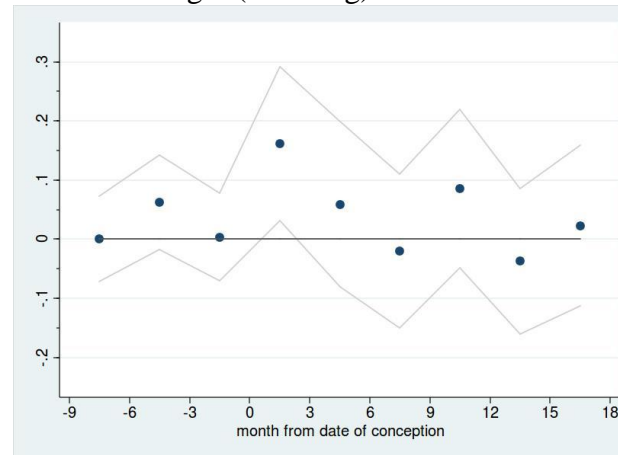


**Figure 11: Effect on Outcomes by Trimester since Conception**

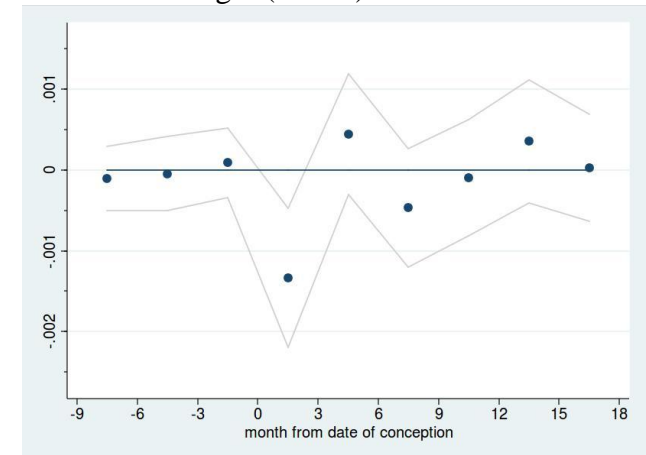
Birthweight (g.)



Low birthweight (<2.500 g)



Gestational length (weeks)



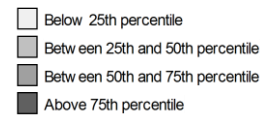
Note: The picture reports estimated effects of local homicide rate at different points before, during and after pregnancy. Trimesters are expressed since (from) the date of conception denoted by 0.

**Figure 12: Small Municipality Characteristics**

Fraction low-weight births



Homicide rate (in the street)



Note: The pictures report, respectively, the average fraction low-weight births (<2.5 kg) and the homicide rate in public places for municipalities of average size no greater than 5,000 inhabitants.

## References

- Adams-Byers, J., Whitsell, S. S. and Moon, S.M. 2004. "Academic and Social/Emotional Effects of Homogeneous and Heterogeneous Grouping." *Gifted Child Quarterly*, 48(7): 7-20.
- Aizer, A. 2011. "Poverty, Violence and Health: The Impact of Domestic Violence during Pregnancy on Newborn Health." *Journal of Human Resources* 46(3): 518-538.
- Aizer, A., Stroud, L. and Buka, S. 2009. "Poverty, Stress and the Intergenerational Transmission of Human Capital." *Unpublished Manuscript*.
- Amarante, V., Manacorda, M., Miguel, E. and Vigorito, A. 2011. "Do Cash Transfers Improve Birth Outcomes? Evidence from Matched Vital Statistics, Social Security and Program Data." *NBER Working Paper 17690*. Cambridge, United States: National Bureau of Economic Research.
- Alderman, H. and Behrman, J. 2006. "Reducing the Incidence of Low Birthweight in Low-Income Countries Has Substantial Economic Benefits." *World Bank Research Observer* 21(1): 25-48.
- Almond, D. 2006. "Is the 1918 Influenza Pandemic Over? Long-term Effects of *In Utero* Influenza Exposure in the Post-1940 U.S. Population." *Journal of Political Economy* 114(4): 672-712.
- Almond, D., Chay, K. Y. and Lee, D. S. 2005. "The Costs of Low Birthweight." *The Quarterly Journal of Economics* 120(3): 1031-1083.
- Almond, D. and Currie, J. 2011a. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives* 25(3): 153-172.
- , 2011b. "Human Capital Development before Age Five." In: O. Ashenfelter and D. Card, Editors. *Handbook of Labor Economics* 4: 1315-1486.

- Almond, D., Edlund, L. and Palme, M. 2009. "Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden." *Quarterly Journal of Economics* 124(4): 1729-1772.
- Almond, D. and Mazumder, B. 2011. "Health Capital and the Prenatal Environment: The Effect of Ramadan Observance during Pregnancy." *American Economic Journal: Applied Economics* 3(4): 56-85.
- Almond, D., Hoynes, H. W. and Whitmore Schanzenbach, D. 2011. "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes." *The Review of Economics and Statistics* 93(2): 387-403.
- Ammermueller, A. and Pischke, J.-S. 2009. "Peer Effects in European Primary Schools: Evidence From the Progress in International Reading Literacy Study." *Journal of Labor Economics*, 27(3): 315-348.
- Angrist, J., Bettinger, E., Bloom, E., King, E. and Kremer, M., 2002. "Vouchers for Private Schooling in Colombia: Evidence from Randomized Natural Experiments." *American Economic Review*, 92(5): 1535-1558.
- Angrist, J., Bettinger, E., and Kremer, M., 2006. "Long-term Consequences of Secondary School Vouchers: Evidence from Administrative Records in Columbia." *American Economic Review*, 96(3): 847-862.
- Angrist, J. and Krueger, A.B. 1991. "Does Compulsory School Attendance Affect Education and Earnings?" *Quarterly Journal of Economics*, 106(4): 979-1014.
- Angrist, J. and Lang, K. 2004. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *American Economic Review*, 94(5): 1613-1634.
- Angrist, J. and Lavy, V. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Children's Academic Achievement." *Quarterly Journal of Economics*, 114(2): 533-575.

- Ashenfelter, O., 1978. Estimating the Effect of Training Programs on Earnings. *Review of Economics and Statistics*, 60(1): 47-57.
- Athey, S., Imbens, G. 2006. Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica*, 74(2): 431-497.
- Banerjee, A., Duflo, E., Postel-Vinay, G. and Watts, T. 2010. "Long-run Health Effects of Income Shocks: Wine and Phylloxera." *Review of Economics and Statistics* 92(4): 714-728.
- Bandiera, O., Barankay, I. and Rasul, I. 2010. "Social Incentives in the Workplace." *Review of Economic Studies*, 77(2): 417-459.
- Barreca, A. 2010. "The Long-Term Economic Impact of In Utero and Postnatal Exposure to Malaria." *Journal of Human Resources* 45(4): 865-892.
- Bayer, P., Hjalmarsson, R. and Pozen, D. 2009. "Building Criminal Capital Behind Bars: Peer Effects in Juvenile Corrections." *Quarterly Journal of Economics*, 124(1): 105-147.
- Betts, J.R. and Shkolnik, J.L. 1999. "The Effects of Ability Grouping on Student Achievement and Resource Allocation in Secondary Schools." *Economics of Education Review*, 19(1): 1-15.
- Camacho, A. 2008. "Stress and Birthweight: Evidence from Terrorist Attacks." *American Economic Review* 98(2): 511-515.
- Carrell, S., Fullerton, R. and West, J. 2009. "Does Your Cohort Matter? Measuring Peer Effects in College Achievement." *Journal of Labour Economics* 27(3): 439-464.
- Carrell, S., Sacerdote, B., and West, J. 2013. "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation." *Econometrica* 81(3): 855-882.
- Cascio, E. and Whitmore Schanzenbach, D. 2007. "First in the Class? Age and the Education Production Function." *NBER Working Paper* 13663.

- Chay, K. and Greenstone, M. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *Quarterly Journal of Economics* 118(3): 1121-1167.
- Couttolene, B., Cano, I., Carneiro, P. and Phebo, L. 2000. "Violencia y Policía en Rio de Janeiro." In: J. L. Londoño, A. Gaviria, and R. Guerrero (Eds.). *Asalto al Desarrollo: Violencia en América Latina*. Washington, DC, United States: Inter-American Development Bank.
- Currie, J. 2011. "Inequality at Birth: Some Causes and Consequences." *American Economic Review* 101(3): 1-22.
- Currie, J. and Moretti, E. 2007. "Biology as Destiny? Short- and Long-term Determinants of Intergenerational Transmission of Birthweight." *Journal of Labor Economics* 25(2): 231-264.
- Currie, J. and Walker, R. 2011. "Traffic Congestion and Infant Health: Evidence from E-ZPass." *American Economic Journals: Applied* 3(1): 65-90.
- De Giorgi, G., Pellizzari, M. and Woolston, W.G. 2012. "Class Size and Class Heterogeneity." *Journal of the European Economic Association*, 10(4): 795-830.
- De Janvry, A., Finan, F. and Sadoulet, E., 2006. "Evaluating Brazil's Bolsa Escola Program: Impact on Schooling and Municipal Roles." *Unpublished manuscript*.
- De Mello, L., Hoppe, M., 2005. Education Attainment in Brazil: the Experience of FUNDEF. *OECD Economics Department Working Papers* 424.
- Donald, S., Lang, K., 2007. "Inference with Differences-in-Differences and Other Panel Data." *Review of Economics and Statistics*, 89(2): 221-233.
- Dong, Y., 2009. "Kept Back to Get Ahead? Kindergarten Retention and Academic Performance." *European Economic Review*, 54(2): 219-236.

- Duflo, E., Dupas, P. and Kremer, M. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review*, 101(5): 1739–1774.
- Ecclestone, M. 2012. "In Utero Exposure to Maternal Stress: Effects of 9/11 on Birth and Early Schooling Outcomes in New York City." *Unpublished manuscript*. Cambridge, United States: Harvard University.
- Eide, E. and Showalter, M., 2001. "The Effect of Grade Retention on Educational and Labor Market Outcomes." *Economics of Education Review*, 20(6): 563-576.
- Figlio, D. and Page, M. 2002. "School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality?" *Journal of Urban Economics*, 51(3): 497-514.
- Firpo, S., Fortin, M., Lemieux, T. 2009. Unconditional Quantile Regressions. *Econometrica*, 77(3), 953-973.
- FUNASA. 2001. "*Manual de Procedimentos do Sistema de Informações sobre Nascidos Vivos*." Ministério da Saúde. Brasília, Brazil: Fundação Nacional de Saúde.
- Gamper-Rabindran, S., Khan, S., Timmins, C. 2010. The Impact of Piped Water Provision on Infant Mortality in Brazil: A Quantile Panel Data Approach. *Journal of Development Economics*, 92(2), 188-200.
- Gibbons, S. and Telhaj, S. 2012. "Peer Effects: Evidence from Secondary School Transition in England." *IZA Discussion Paper No. 6455*.
- Glick, P. and Sahn, D., 2010. "Early Academic Performance, Grade Repetition, and School Attainment in Senegal: A Panel Data Analysis." *The World Bank Economic Review*, 24(1): 93-120.

- Gomes-Neto, J.B. and Hanushek, E., 1994. "Causes and Consequences of Grade Retention: Evidence from Brazil." *Economic Development and Cultural Change*, 43(1): 117-148.
- Havnes, T., Mogstad, M. 2010. Is Universal Child Care Levelling the Playing Field? Evidence from Non-Linear Difference-in-differences. *IZA Discussion Paper No.* 4978.
- Hanushek, E., Kain, J.F., Markman, .M. and Rivkin. S.G. 2003. "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics*, 18(5): 527-544.
- Heinemann, A. and Verner, D. 2006. "Crime and Violence in Development: A Literature Review of Latin America and the Caribbean." *World Bank Policy Research Working Paper 4041*. Washington, DC, United States: World Bank.
- Holmes, T.C., 1989. "Grade Level Retention Effects: A Meta-Analysis of Research Studies." In: *Flunking Grades: Research and Policies on Retention*, ed. Lorrie A. Shepard and Mary L. Smith, 16-33. London: Falmer Press.
- Hoxby, C.M. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." *NBER Working Paper* 7867.
- Hoxby, C.M. and Weingarth, G. 2006. "Taking Race out of the Equation: School Reassignment and the Structure of Peer Effects." *Unpublished Manuscript*.
- IBGE (Brazilian Institute of Geography and Statistics), 2007. "*Contas Nacionais No. 21: Contas Regionais do Brasil 2002-2005*." Rio de Janeiro.
- Imbens, G.W. and Lemieux, T. 2007. "Regression Discontinuity Designs: a Guide to Practice." *Journal of Econometrics*, 142(2): 615-635.
- Imbens, G. and Wooldridge, J., 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economics Literature*, 47(1): 5-86.
- INEP (National Institute for Educational Studies and Research). 2007. "*IDEB - Prova Brasil/Saeb 2007*." Brasilia.



- Jacob, B.A. 2005. Accountability, Incentives and Behaviour: Evidence from School Reform in Chicago. *Journal of Public Economics*, 89(5-6), 761-796.
- Jacob, B. and Lefgren, L., 2004. "Remedial Education and Student Achievement: a Regression-Discontinuity Analysis." *Review of Economics and Statistics*, 86(1): 226-244.
- Jacob, B. and Lefgren, L., 2009. "The Effect of Grade Retention on High School Completion." *American Economic Journal: Applied Economics*, 1(3): 33-58.
- Kahn, T. 1999. "Os Custos Da Violência: Quanto se Gasta ou Deixa de Ganhar por Causa do Crime no Estado de São Paulo." *São Paulo em Perspectiva* 13(4): 42-48.
- Kelly, E. 2011. "The Scourge of Asian Flu: In Utero Exposure to Pandemic Influenza and the Development of a Cohort of British Children." *Journal of Human Resources* 46(4): 669-694.
- Kodde, D.A., Palm, F.C. and Pfann, G.A. 1990. "Asymptotic Least-squares Estimation Efficiency Considerations and Applications." *Journal of Applied Econometrics*, 5(3): 229-243.
- Koenker, R. 2004. Quantile Regression for Longitudinal Data. *Journal of Multivariate Analysis*, 91(1), 74-89.
- Kremer, M. 1997. "How Much does Sorting Increase Inequality." *Quarterly Journal of Economics*, 112(1): 115-139.
- Kremer, M., Miguel, E. and Thornton, R., 2009. "Incentives to Learn." *Review of Economics and Statistics*, 91(3): 437-456.
- Lam, D. and Marteleto, L. 2006. Small Families and Large Cohorts: The Impact of the Demographic Transition on Schooling in Brazil. In: Lloyd, C., Behrman, J., Stromquist, N., and Cohen, B. (eds). *The Changing Transition to Adulthood in Developing Countries*. The National Academic Press, Washington.

- Lamarche, C. 2011. Measuring the Incentives to Learn in Colombia Using New Quantile Regression Approaches. *Journal of Development Economics*, 96(2), 278-288.
- Lavy, V., Paserman, D. and Schlosser, A. 2012. "Inside the Black Box of Ability Peer Effects: Evidence from Variation in Low Achievers in the Classroom." *Economic Journal*, 122(559): 208-237.
- Lavy, V. and Schlosser, A. 2011. "Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics*, 3(2): 1-33.
- Lavy, V., Silva, O. and Weinhardt, F. 2012. "The Good, the Bad and the Average: Evidence on the Scale and Nature of Ability Peer Effects in Schools." *Journal of Labour Economics*, 30(2): 367-414.
- Lazear, E. 2001. "Educational Production." *Quarterly Journal of Economics*, 116(3): 777-803.
- Lee, D. and Lemieux, T. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2): 281-355.
- Lindert, K., Linder, A., Hobbs, J. and de la Brière, B. 2007. "The Nuts and Bolts of Brazil's Bolsa Família Program: Implementing Conditional Cash Transfers in a Decentralized Context." *Social Protection Discussion Paper 0709*, World Bank.
- Lyle, D.S. 2009. "The Effects of Peer Group Heterogeneity on the Production of Human Capital at West Point." *American Economic Journal: Applied Economics*, 1(4): 69-84.
- Mainardes, J., 2010. "Moving Away from a Graded System: a Policy Analysis of the Primary Education Organized in Cycles in Brazil." Lampert, Saarbrücken.
- Manacorda, M., 2012. "The Cost of Grade Retention." *Review of Economics and Statistics*, 94(2): 596-606.

- Mansour, H. and Rees, D. 2012. "Armed Conflict and Birthweight: Evidence from the al-Aqsa Intifada." *Journal of Development Economics* 99(1): 190-99.
- Mas, A. and Moretti, E. 2009. "Peers at Work." *American Economic Review*, 99(1): 112-145.
- Manski, C.F. 1993. "Identification of Endogenous Social Effects: the Reflection Problem." *Review of Economic Studies*, 60(3): 531-542.
- McCrary, J. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: a Density Test." *Journal of Econometrics*, 142(2): 698-714.
- Ministry of Education (MEC), 1996. *Lei de Diretrizes e Bases da Educação Nacional*, No. 9394, Art. 24, December 1996, Brasilia.
- Ministry of Education (MEC). 2004. "*Ensino Fundamental de Nove Anos – Orientações Gerais*." Secretariat of Basic Education, Federal Brazilian Ministry of Education. Brasília.
- Nilsson, P. 2009. "The Long-term Effects of Early Childhood Lead Exposure: Evidence from Sharp Changes in Local Air Lead Levels Induced by the Phase-out of Leaded Gasoline." *Unpublished manuscript*. Uppsala, Sweden: Uppsala University.
- OECD, 2005. "*OECD Economic Survey of Brazil 2005*." Organization for Cooperation and Development, Paris.
- Ohinata, A. and van Ours, J. 2013. "How Immigrant Children Affect the Academic Achievement of Native Dutch Children." *Economic Journal*, 123(570): 308-331.
- Patrinos, H.A. and Psacharopoulos, G. 1996. "Socioeconomic and Ethnic Determinants of Age-grade Distortion in Bolivian and Guatemalan Primary Schools." *International Journal of Educational Development*, 16(1): 3-14.
- Pino, I.R. and Koslinki, M.C., 1999. "Government Programs to Eliminate Repetition, School Dropout and Exclusion in Brazil." In: *Schooling for Success: Preventing*

- Repetition and Dropout in Latin American Primary Schools*, ed. Laura Randall and Joan B. Anderson. New York: Sharpe.
- Reichenheim, M. E., Ramos de Souza, E., Leite Moraes, C., de Mello Jorge, M., Passos da Silva, C. and Minayo, M. 2011. "Violence and Injuries in Brazil: the Effect, Progress Made, and Challenges Ahead." *Lancet* 377(9781): 1962-1975.
- Reyes, W. 2007. "The Impact of Prenatal Lead Exposure on Infant Health." *NBER Working Paper 13097*. Cambridge, United States: National Bureau of Economic Research.
- Robinson, J.P. 2008. "Evidence of a Differential Effect of Ability Grouping on the Reading Achievement Growth of Language-minority Hispanics." *Educational Evaluation and Policy Analysis*, 30(2): 141–180.
- Rocha, R. and Soares, R. 2012. "Water Scarcity and Birth Outcomes in the Brazilian Semiarid." *IZA Discussion Paper* No. 6773.
- Royer, H. 2009. "Separated at Girth: U.S. Twin Estimates of the Effects of Birthweight." *American Economic Journal: Applied Economics* 1(1): 49-85.
- Sacerdote, B. 2003. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics*, 116(2): 681-704.
- Soares, R. 2010. "Welfare Cost of Crime and Common Violence: A Critical Review." *Discussion Paper 581*, Department of Economics. Rio de Janeiro, Brazil: PUC, Rio.
- Torche, F. 2011. "The Effect of Maternal Stress on Birth Outcomes: Exploiting a Natural Experiment." *Demography* 48(4): 1473-1491.
- UNESCO, 2008. "*EFA Global Monitoring Report*." United Nations Educational, Scientific and Cultural Organization, Paris.
- UNICEF. 2006. "*Low Birthweight: Country, Regional and Global Estimates*." Geneva, Switzerland: United Nations Children's Fund.

- UNODC. 2005. “*Country Report Brazil 2005.*” Regional Office Brazil. Brasilia, Brazil: United Nations Office on Drugs and Crime.
- UNODC. 2011. “*2011 Global Study on Homicide. Trends, Contexts, Data.*” Vienna, Austria: United Nations Office on Drugs and Crime.
- Urquiola, M. 2006. “Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia.” *Review of Economics and Statistics*, 88(1): 171–177.
- Van der Klaauw, W. 2002. “Estimating the Effect of Financial Aid Offers on College Enrolment: A Regression-discontinuity Approach.” *International Economic Review*, 43(4): 1249-1287.
- Velasco Rondon, V. and Andrade, M. 2003. “Custos da Criminalidade em Belo Horizonte.” *Economía* 4(2): 223-259.
- Victora, C., Kirkwood, B., Ashworth, A., Black, R., Rogers, S., Sazawal, S., Campbell, H. and Gore, S. 1999. “Potential Interventions for the Prevention of Childhood Pneumonia in Developing Countries: Improving Nutrition.” *American Journal of Clinical Nutrition* 70(3): 309-320.
- Whitmore, D. 2005. “Resource and Peer Impacts on Girls” Academic Achievement: Evidence from a Randomized Experiment.’ *American Economic Review*, 95(2): 199–203.
- Wolfowitz, J. 1957. “The Minimum Distance Method.” *The Annals of Mathematical Statistics*, 28(1): 75-88.
- Zimmer, R. 2003. “A New Twist in the Educational Tracking Debate.” *Economics of Education Review*, 22(3): 307-315.
- Zimmerman, D.J. 2003. “Peer Effects in Academic Outcomes: Evidence from a Natural Experiment.” *Review of Economics and Statistics*, 85(1): 9-23.